The Molecular Signature for Local Adaptation in the Seagrass *Posidonia oceanica*

Thesis

Daniela D'Esposito

Degree in Biology

University Federico II of Naples, Italy

The Molecular Signature for Local Adaptation in the Seagrass *Posidonia oceanica*

Doctor of Philosophy

The Open University of London

ARC: Stazione Zoologica Anton Dohrn of Naples, Italy

DATE of SUBMISSION : 31 MAY 2013
DATE of AWARD : 15 SEPTEMBER 2013

May 2013

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Thesis abstract

In the last century, seagrass ecosystems have suffered a worldwide decline ascribed to multiple environmental stressors, among which the reduction of light available for the photosynthesis and the increase in temperature represent the strongest constraints for their growth and survival. Despite conservation, this decline is at present still continuing.

In order to understand the genetic adaptive response to light and temperature in the seagrass *Posidonia oceanica,* two different strategies have been pursued: a genome scan approach along a latitudinal and a bathymetric gradient and a differential gene expression analysis along the bathymetric gradient, where light and temperature were the two main selective factors.

For the genome scan approach 6 populations (Delimara - Malta, Lacco Ameno – Island of Ischia, Italy, Marettimo Island- Italy, Meloria - Italy, Piombino - Italy and Stareso – Corsica, France) were sampled along the bathymetric gradient at two different depths (-5m and -25m). The same populations were used for the latitudinal gradient analysis by grouping them on the basis of their geographic location (Southern group: Delimara, Lacco Ameno and Marettimo; Northern group: Meloria, Piombino and Stareso).

No genes under selection were identified in the genome scan along the bathymetric gradient. Three putative genes under selection were identified in the genome scan along the latitudinal gradient and were involved in the photosynthesis and in the translation process.

For assessing differential gene expression, a transcriptome sequencing of plants sampled at two different depths and different times of the day in the Stareso meadow was performed by RNAseq technology. The analysis highlighted the capability of plants living in shallow waters to cope with environmental stresses imposed by high light and high temperature. Transcriptome data generated from this study increased the resources available in *P. oceanica* and will be very useful for further investigations of the adaptation of in this plant.

Thesis overview: research background, objectives and outline of the thesis

The unrelenting threat of seagrass loss mainly resulting from reduced light availability and from increased temperature, as result of human activities, has motivated this study aimed to elucidate the molecular mechanisms responsible for the adaptation of the seagrass *P. oceanica* to environmental gradients, mainly consisting of different regimes of light and temperature, two limiting factors for the growth and the survival of this species. This topic is dealt with approaches both at genomic and transcriptome level. Adaptation by genetic change may, in fact, involve structural changes within the coding regions of the candidate gene and regulatory changes that alter the transcript abundance. In particular, the objectives of this study are:

1) the detection of candidate genes under selection related to light and temperature changes, by a genome scan approach based on microsatellites and performed along a latitudinal and bathymetric gradient;

2) the enrichment of genomic resource in *P. oceanica* and the identification of biological processes whose expression pattern is affected in different condition of light and temperature, by means of RNAseq technology. Following is reported the structure of the thesis. *Chapter 1* gives an overview of the seagrasses ecosystems, ecological functions and threats, with a particular focus on *P. oceanica*. A background of genetic studies in *P. oceanica* is also presented. To reach the first objective of the thesis primers for EST-linked microsatellites (E-SSRs) were developed in *P. oceanica*. *Chapter 2* presents a technical note of the development of

primers for 51 novel EST-SSRs, as well as the their assembling in multiplex PCR. A new assembly in multiplex PCR is also showed for the 13 anonymous microsatellites (A-SSRs) already available in *Posidonia*. In addition, this chapter reports the analysis of the genetic diversity in two *P. oceanica* meadows performed with the total set of 64 microsatellites. Because A-SSR and E-SSRs come from different regions of the genome and they could have different properties in the assessment of genetic diversity a study of the performance of these two classes of markers has also been carried out that is presented in *Chapter 3*. In this chapter is reported the comparison of the two classes of markers in relation to the assessment of clonal diversity, summary statistics and in the detection of genetic structure in a survey involving 6 populations. *Chapter 4* describes the genome scan approach to identify candidate genes related to light and temperature. The genome scan is presented in parallel for the latitudinal gradient, involving populations belonging to a northern group and a southern group in the Mediterranean Sea, and for the bathymetric gradient, involving populations sampled at two depths, -5 m and -25 m. To reach the second objective, Next-Generation Sequencing (NGS) technology is applied to sequence the transcriptome of samples collected at different depths and in different times of the day. In *Chapter 5* are presented the results obtained from the massive parallel sequencing (RNAseq) and the enrichment analysis to identify the biological processes, the cell component and the molecular functions over-represented in particular conditions of light and temperature. *Chapter 6* concludes the thesis by discussing the overall contribution of the research. In addition, chapter 6 discusses the possible direction of future research.

Data from chapter 2 have been published in the paper: *Arranz S.E., Avarre J.C., Balasundaram C., et al. 2013 Permanent Genetic Resources added to Molecular Ecology Resources Database 1 December 2012–31 January 2013 2013 (D'Esposito D., Orsini L., Procaccini G. Development of 51 novel EST-linked microsatellites in the Mediterranean seagrass Posidonia oceanica). Molecular Ecology Resources, 13 (3): 546-549.*

Preliminary data from chapter 3 have been presented in the paper: *D'Esposito D, Dattolo E, Badalamenti F, Orsini L, Procaccini G (2012) Comparative analysis of genetic diversity of Posidonia oceanica along a depth gradient using neutral and selective/non neutral microsatellite markers. Biologia Marina Mediterranea 19 (1): 45-48*

Conclusive data from chapter 3 and data from the other two chapters will be included in three independent publications.

List of Figures

Chapter 1

Chapter 2

Chapter 3

Appendix Chapter 5

List of Tables

Chapter 2

Chapter 3

Appendix Chapter 3

Chapter 4

Appendix Chapter 4

Chapter 5

Glossary

A

-Adaptation: heritable changes in genotype or phenotype that result in increased fitness.

-Admixture: the proportion of new genetic combinations in hybrid populations through recombination.

-Allele: one of a pair, or series, of alternative forms of a gene that occurs at a given locus in a chromosome.

-Allele frequency: the proportion of one allele relative to all alleles at a locus in a population.

-Allelic richness: standardize measure of the number of alleles per locus independent of the sample size.

-Analysis of Molecular Variance (AMOVA): method for studying molecular variation within a population or species.

B

-Balancing selection: a general type of selective regime that favours the maintenance of diversity in a population. There are two main mechanisms by

which balancing selection preserves polymorphism: heterozygote advantage (or overdominance) and frequency-dependent selection.

-**Bayesian:** a framework of statistical inference in which previous beliefs (or data) and likelihoods are combined to estimate a parameter of interest given the observed data.

C

-**Climate change:** a statistically significant variation in either the mean state of the climate or in its variability, persisting for an extended period (typically decades or longer). Climate change may be due to natural internal processes or external forcings, or to persistent anthropogenic changes in the composition of the atmosphere or in land use.

-**Clonality:** life-history strategy allowing organisms to produce offspring without sexual reproduction, hence typically genetically identical, at the exception of possible somatic mutations, to themselves.

-**Clonal reproduction:** asexual reproduction through propagation of plants parts not involving seeds, also called clonal growth or vegetative reproduction.

-**Cluster:** a genetic grouping of individuals, often from several nearby sampling locations, that shares a similar multilocus allele-frequency distribution.

-**Coding region:** portion of a gene that encodes a protein.

-**Contig:** an abbreviation for contiguous sequence; used to indicate a contiguous piece of DNA that is assembled from short overlapping sequences.

D

-*De novo* **transcriptome assembly:** assembly strategy that does not use a reference genome. It leverages the redundancy of short-read sequencing to find overlaps between the reads and assemble them into transcripts.

E

-**Ecosystem:** a biotic community and its abiotic environment.

-**Environmental factor:** any factor, abiotic or biotic, that influences living organisms; also known as an ecological factor or ecofactor.

-**Expressed Sequence Tags (ESTs):** short DNA sequences (several hundred base pairs) produced by reverse transcription of mRNA into DNA.

F

-**Fitness:** a measure of the capacity of an organism to survive and reproduce.

-**Fixation:** the increase in the frequency of a genetic variant in a population to 100%.

-F_{ST}: index of genetic divergence between populations. A standardized measure of the distribution of genetic variation between populations on a scale between 0 (identical allele frequencies in populations) and 1 (populations fixed for different alleles).

-**Gene flow**: the exchange of genes between populations that is caused by the dispersal of propagules or individuals.

-**Gene Ontology (GO):** a widely used classification system of gene functions and other gene attributes that uses a controlled vocabulary. The ontology covers three domains; cellular components, molecular functions and.

-**Gene Ontology Enrichment analysis:** statistical framework to find molecular functions, biological processes and cellular locations significantly associated with gene groups. To test for association each gene is categorized in two ways: first whether it is annotated with the term under consideration or not, and second, whether it belongs to the test group or not. Using Fisher's exact tests *p*-values are computed that allow to detect and quantify association between the two categorizations.

-**Genet**: a genetic individual composed of multiply connected or physiologically autonomous ramets (vegetative or reproductive shoots).

-**Genetic drift:** random fluctuations in allele frequencies between generations owing to sampling effects.

-**Genetic diversity:** measure of genetic variability based on gene frequencies in the population.

-**Genetic hitch-hiking:** the process by which a neutral, or in some cases deleterious, mutation may increase in population frequency owing to linkage with a positively selected mutation. (see also Positive selection).

-**Genome:** the entire complement of genetic material in an organism.

-**Genome scan:** a survey of genetic variability across a large number of loci in a large number of individuals.

-**Genotype:** the combination of alleles present at a locus in an individual.

H

-**Hardy-Weinberg equilibrium:** the stable frequency distribution of genotypes, AA, Aa, and aa, in the proportions p^2, 2pq, and q^2, respectively (where p and q are the frequencies of the alleles, A and a), that is a consequence of random mating in the absence of mutation, migration, natural selection or random drift.

-**Heterozygosity:** the proportion of individuals in a population that carry two different alleles at a locus.

-**Heterozygote:** a diploid individual that has different alleles at one or more genetic loci (cf. Homozygote)

-**Homozygote:** a diploid individual that has identical alleles at one or more genetic loci (cf. Heterozygote).

L

Linkage disequilibrium: the correlation between alleles across loci.

-Locus (pl. loci): the specific place on a chromosome where a gene or particular piece of DNA is located.

M

-Marker: an identifiable physical location on a chromosome whose inheritance can be monitored (e.g. gene, restriction enzyme site or RFLP marker).

-Microsatellite: a 2-7 base pair repeating motif (allele vary in length); synonyms include simple sequence repeat (SSR) and short tandem repeat (STR).

-Migration: movement of individuals between otherwise reproductively isolated populations.

-Monomorphic locus: locus in a population that has only one allele. All individuals are homozygous for the same allele. Lacking of genetic diversity. (cf. Polymorphic locus).

-Multiplex PCR: co-amplification of multiple microsatellites in a single PCR reaction.

-Mutation: change in a DNA sequence.

N

-Natural selection: the process in nature by which, according to Darwin's theory of evolution, only the organisms best adapted to their environment tend to survive

and transmit their genetic characteristics in increasing numbers to succeeding generations while those less adapted tend to be eliminated.

-**Neutral distribution or null distribution:** the distribution (or range) of values across which we expect to observe the value of the test statistic if the null hypothesis is true (for example, neutrality). When conducting a standard t-test, t is the test statistic and the null distribution is the normal (Gaussian) distribution with t degrees of freedom.

-**Neutral loci:** loci that are not evolving directly in response to selection, the dynamics of which are controlled mainly by genetic drift and migration. These loci can, however, be influenced by selection on nearby (linked) loci.

-**Null alleles:** non-amplifying alleles that result in an apparent homozygote when present in heterozygote state and in the lack of amplification when present in homozygote state. In the latter case, they can be confounded with reaction failure. Null alleles are produced by mutations in the flanking region, at primer binding sites.

O

-**Outlier loci:** genome locations (or markers or base pairs) that show behaviour or patterns of variation that are extremely divergent from the rest of the genome (locus-specific effects), as revealed by simulations or statistical tests (see also Outlier tests).

-Outlier tests: popular approaches for detecting selection that identify loci that present extreme values for a given statistic from empirical genome-wide genetic data. It is assumed that these 'outlier loci' are enriched for loci targeted by selection (see also Outlier loci).

P

-Paired-end sequencing technology: library construction and sequencing strategy in which both ends of a DNA fragment are sequenced to produce pairs of reads (mate pairs).

-Polymerase Chain Reaction (PCR): a procedure that produces millions of copies of a short segment of DNA through repeated cycles of: 1) denaturation, 2) annealing, and 3) elongation.

-Phenotype: the observable physical and/or biochemical characteristics of the expression of a gene.

-Phenotypic plasticity: the range of phenotypes a single genotype can express as a function of its environment.

-Polymorphic Information Content (PIC): refers to a value of a marker for detecting polymorphism within a population. It depends on the number of detectable alleles and the distribution of their frequencies.

-Polymorphic locus: locus in a population that has two or more alleles. Polymorphic loci are usually defined as having the most frequent allele at a

frequency of less than 0.99, or less than 0.95 (to minimize problems with different sample sizes). (cf. Monomorphic locus).

-**Population:** a group of individuals who share a common genepool and have the potential to interbreed.

-**Population genetics:** the quantitative study and measurement of populations in statistical terms, e.g. the study of genetic phenomena in terms of standard statistical parameters such as frequency tables and distributions, means, variance and standard deviations.

-**Positive selection:** also known as directional or Darwinian selection, this refers to selection acting upon newly arisen (or previously rare) advantageous mutations. When an advantageous mutation increases in frequency in the population as a result of positive selection, linked neutral variation will be dragged along with it – a process known as genetic hitch-hiking. As a consequence, variation that is not associated with the selected allele is eliminated, resulting in a selective sweep that leads to an overall reduction of genetic diversity around the selected site.

-**Principal Coordinate Analysis (PCoA):** multivariate technique that allows to find and plot the major patterns within a multivariate dataset e.g. multiple loci and multiple samples.

R

-**Ramet:** the unit of clonal growth in plants; corresponding to a shoot that can be vegetative or reproductive.

-RNA-seq or RNA sequencing: experimental protocol that uses next generation sequencing technologies to sequence the RNA molecules within a biological sample in an effort to determine the primary sequence and relative abundance of each RNA.

S

-Selective sweep: a reduction in genetic variation in DNA that surrounds a locus that is under selection.

-Signaling cascades: cascades of events that mediate cellular responses to external signals.

-Signature of selection: the molecular footprint of a selection event from the recent past (for example, an excess of rare alleles at a locus relative to the abundance of rare alleles at loci across the rest of the genome).

-Split peaks: PCR artefacts caused by the non-template addition of a nucleotide (generally an adenine) to PCR fragments by the Taq polymerase. When this adenylation is incomplete, it results in double peaks (the original fragment and an additional peak 1 bp longer corresponding to the adenylated fragment), thereby compromising automatic peak recognition, particularly for heterozygote genotypes with nearby alleles.

-SSR: see Microsatellite.

-STR: see Microsatellite.

-Summary statistic: a parameter estimate (such as F_{ST} or F_{IS}) that quantifies attributes of the data sampled from a population of interest.

T

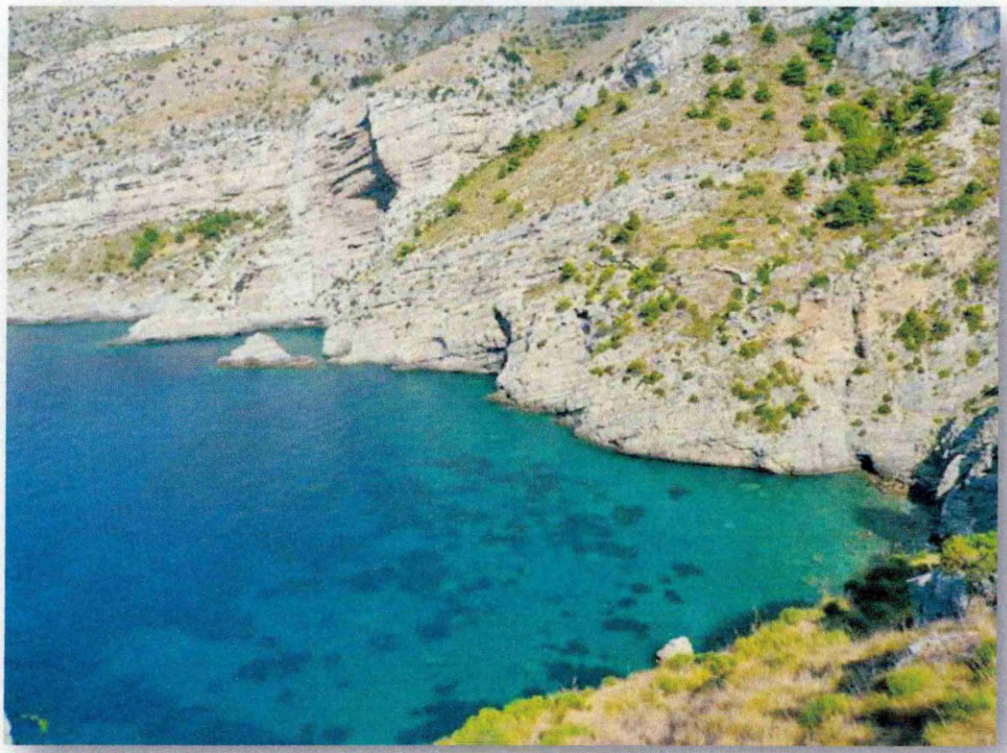
-Type I error: statistical error in which a true null hypothesis is rejected.

-Transcriptome: the transcriptome is the complete set of transcripts in a cell, both in terms of type and quantity.

-Transcriptome assembly: reconstruction of full-length transcripts starting from the short sequences (reads) generated by the sequencing.

Table of Contents

Chapter 1



1-Introduction

1.1-Seagrasses

Seagrasses are monocotyledonous angiosperms that live completely submerged in marine and estuarine environments. Seagrasses belong to the order of Alismatales which includes 11 families of aquatic-freshwater species and 4 families that are fully marine (Wissler et al., 2011). The marine families include the Posidoniaceae, Zosteraceae, Hydrocharitaceae, and Cymodoceaceae, and have originated in the Cretaceous (den Hartog, 1970). Phylogenetic analysis of members of the entire order, based on the plastid gene encoding for RuBisCO large subunit (Les et al., 1997), indicates that the return into the sea occurred at least three times independently through parallel evolution from a common aquatic-freshwater ancestor of terrestrial origin.

Seagrasses are distributed across the globe (Green and Short, 2003, Figure 1.1) but, dissimilarly to other taxonomic groups with worldwide distribution, they exhibit a low taxonomic diversity with approximately 60 species worldwide, compared with approximately 250,000 terrestrial angiosperms (Orth et al., 2006).

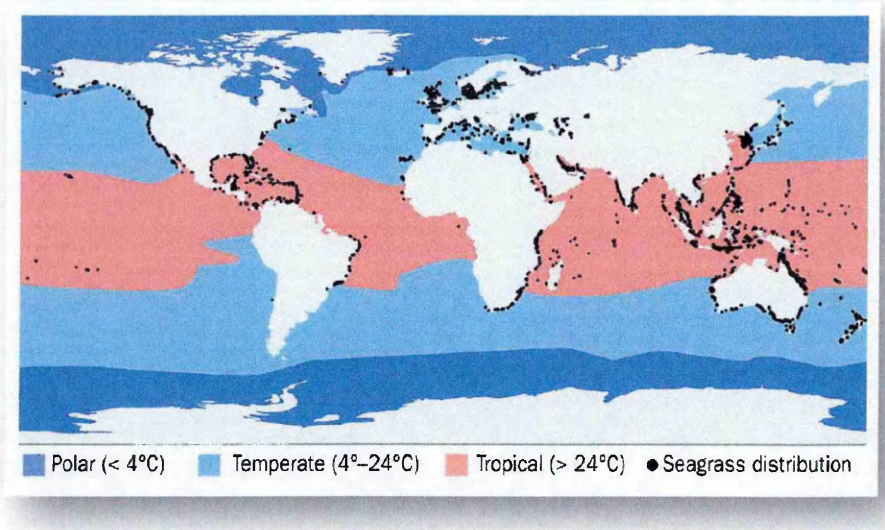


Figure 1.1. Current global distribution of seagrass in relation to mean ocean temperature. Regional divisions are based on polar , temperate and tropical climate (Green and Short 2003).

Seagrasses have similar organs and tissues as the other flowering plants but to enable survival in marine environment they require ecological, physiological and morphological adaptations (den Hartog 1970, Les et al. 1997, Kuo and den Hartog, 2006) that are distinctive from terrestrial angiosperms.

Seagrasses are able to regulate the osmotic pressure of their cell fluids to cope with the high, and often varying, salinity of seawater (Kuo and den Hartog, 2006).

The photosynthetic apparatus needs to be modulated to accommodate the changes in light attenuation through the water depth (Dalla Via et al., 1998). Consequently, the overall light intensity is decreased and the wavelength composition of sunlight reaching underwater plants is different. Accordingly, seagrasses have one of the highest light requirements among angiosperms (Dennison, 1993; Orth et al., 2006). One factor contributing to these high light requirements is the reducing sediments to which seagrasses are rooted. These sediments challenge seagrass root tissues

with anaerobic conditions since marine sediments are often oxygen deficient. When the internal transport of oxygen from shoot to root tissues is not sufficient, seagrasses can be forced to resort to fermentative metabolism (Terrados, 1999; Touchette and Burkholder , 2000).

Submergence also exposes organisms to the forces of wave action and tidal currents that affect reproductive functions and reduce the availability of carbon dioxide (CO_2, Wissler et al., 2011). Consequently, seagrasses have strap shaped leaves and a resistant anchoring system to withstand water movements and they have evolved to propagate via hydrophilous (water-mediated) pollination (Cox, 1993). Hydrophilous pollination may take place thanks to regulation of the floating capacity of pollen, precisely gauged to meet the female inflorescence (flower peduncle, style and stigma) and the viscous, insoluble substances released by the stigma to which pollen sticks. These processes favor germination of the pollen tube and its growth towards the female gamete. Seed maturation and later dispersal are also extraordinary examples of evolutionary adaptations for successful colonization of submerged habitats, with many different types of substrates and seasonal rhythms (Cox, 1993).

In order to carry out photosynthesis and to cope with the reduced availability of carbon dioxide (CO_2) in marine environment, seagrasses, use bicarbonate ions which they obtain from the hydrolysis of carbonic acid (Invers et al., 1999). An active protein, an enzyme called carbonic anhydrase, favors the diffusion of bicarbonate ions through photosynthetic epidermal cells. This enzyme dehydrates carbonic acid, releasing carbon dioxide, which is then reduced to carbon. This

element is contained in several energy-storing molecules to produce glucose, a source of energy for all plant functions.

Leaf morphological adaptation to the aquatic environment also includes an enlarged aerenchyma system, thin cuticle, small epidermal cells with thick walls, concentrations of chloroplast in the epidermal cells and lack of stomata (Kuo and den Hartog, 2006).

1.1.1-Seagrass architecture and growth

Seagrasses share a common architecture, all species being clonal, rhizomatous plants (Hemminga and Duarte, 2000). One of the characteristics of these plants is their modular structure (Figure 1.2) involving the iteration of a base unit (the ramet or module) along a horizontal rhizome. All ramets on a rhizome are genetically identical and comprise the genetic individual (the genet, Harper 1977).

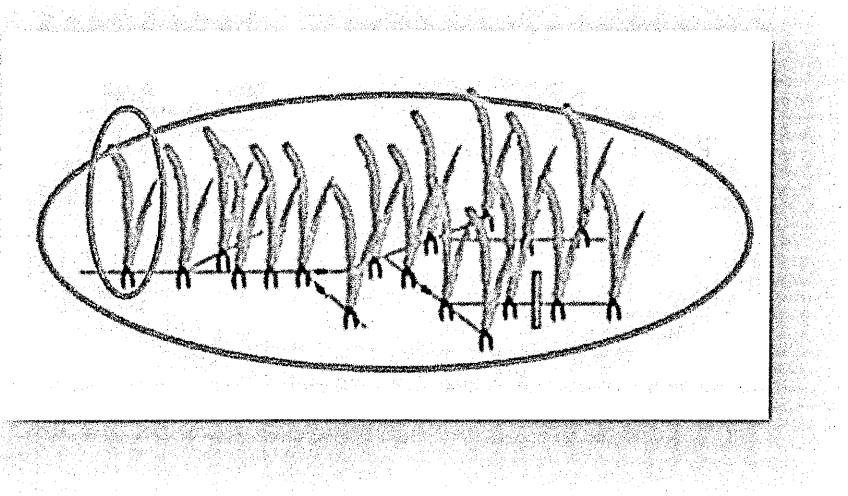


Figure 1.2. Seagrass modular structure. In red is indicated the morphological individual (ramet or module), in green is indicated the genetic individual (genet or clone). Image modified from Procaccini et al. (2007).

Despite the similar architecture of seagrasses, plant size and growth rate vary some orders of magnitude across species. To a large extent, variability in rates between seagrass species reflects differences in plant size, with smaller species growing faster than larger ones (Duarte, 1991). In addition, seagrass growth exhibits substantial plasticity which is considered an important component of their capacity to adapt to environmental change (e.g. Marbà and Duarte, 1994; Marbà et al., 1996; Terrados, 1997; Pérez et al., 1994). Knowledge on seagrass growth rates allows assessment of meadow productivity and seagrass health, as well as forecasting their capacity to survive disturbances.

Vegetative proliferation is the main mechanism of seagrasses to occupy habitat space (Tomlinson, 1974; Duarte and Sand-Jensen, 1990), and thus it is a critical process for seagrass meadows to spread and persist. Most ramets in seagrass populations are produced as rhizomes elongate. Rhizome growth is the process that regulates the rate of formation and the spatial distribution of ramets (and, thus, modules) within seagrass meadows, and, thus, it constrains the development of their populations (Borum et al., 2004). The spread, and maintenance, of seagrass meadows also depend on sexual reproduction since it is the main mechanism regulating patch formation. Hence, information about the effort and success of seagrass reproduction and rhizome growth patterns are essential to predict the time scales of seagrass colonisation and, thus, recovery.

1.1.2-Seagrass ecological functions

Seagrass meadows have an important ecological role in coastal ecosystems and provide high-value ecosystem services compared with other marine and terrestrial habitats (Costanza et al., 1997, Figure 1.3).

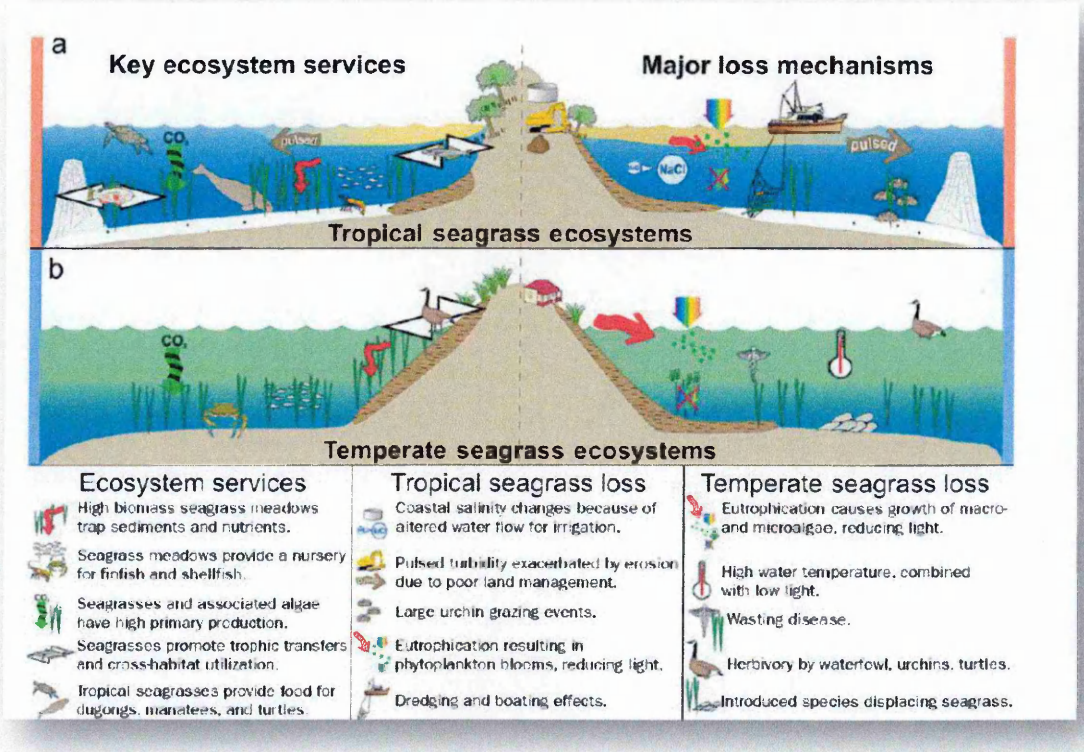


Figure 1.3. Conceptual diagrams for (a) tropical and (b) temperate seagrass ecosystems,detailing key ecosystem services and major mechanisms of seagrass loss (From Orth et al., 2006).

For example, primary production from seagrasses and their associated macro- and microepiphytes rivals or exceeds that of many cultivated terrestrial ecosystems (Duarte and Chiscano, 1999). Seagrasses also provide an enormous source of carbon to the detrital pool, some of which is exported to the deep sea, where it provides a critical supply of organic matter in an extremely food-limited

environment (Suchanek et al., 1985). Much of the excess organic carbon produced is buried within seagrass sediments, which are hotspots for carbon sequestration in the biosphere (Duarte et al., 2005). The structural components of seagrass leaves, rhizomes, and roots modify currents and waves, trapping and storing both sediments and nutrients, and effectively filter nutrient inputs to the coastal ocean (Hemminga and Duarte, 2000). Biodiversity in seagrass meadows is greater than in adjacent unvegetated areas, and faunal densities are orders of magnitude higher inside the meadows (Hemminga and Duarte, 2000). They also serve as a nursery ground, often to juvenile stages of economically important species of finfish and shellfish, although the species vary by region and climate (Beck et al. 2001, Heck et al. 2003). Moreover, seagrasses can be considered as biological sentinels, or "coastal canaries". Changes in seagrass distribution, such as a reduction in the maximum depth limit (Abal and Dennison, 1996) or widespread seagrass loss (Cambridge and McComb, 1984), signal important losses of ecosystem services that seagrasses provide. Seagrasses are sessile, essentially integrating the relevant water quality attributes, such as chlorophyll and turbidity, that affect the light reaching their leaves. Several features of seagrasses and seagrass meadows result in their particular importance in this regard. The widespread distribution of seagrasses throughout tropical and temperate regions (Figure 1.1) allows better assessment of larger-scale trends than do other comparable coastal habitats, such as mangrove, corals, or salt marsh plants, which are limited to only one of these broad geographic regions. Seagrasses also live in shallow, protected coastal waters, directly in the path of watershed nutrient and sediment inputs, and are therefore highly susceptible to these inputs, unlike mangrove forests (which are

largely unaffected by water quality) or coral reefs (which occur farther away from the imputs).

1.1.3-Threats

Synoptic studies to date have examined the distribution, status and trends of seagrass habitat, and have clearly indicated that seagrasses are declining globally (Green and Short, 2003; Orth et al., 2006; Waycott et al., 2009). A synthesis of 215 published studies showed that seagrass habitat disappeared worldwide at a rate of 110 km^2 per year between 1980 and 2006 (Waycott et al., 2009). However, the actual status of individual seagrass species themselves has received little attention. For the first time, the likelihood of extinction of the world's 72 species of seagrasses has been determined under the Categories and Criteria of the International Union for the Conservation of Nature (IUCN) Red List of Threatened Species (Short et al., 2011).

Seagrass losses have been attributed to a broad spectrum of anthropogenic and natural causes (Figure 1.3): Globally, the primary impact to seagrasses is loss of water clarity and quality due to both eutrophication, i.e., phytoplankton and nuisance seaweed blooms (Burkholder et al., 2007), and sediment loading, i.e., suspended sediments and siltation (Dennison et al., 1993; de Boer, 2007). Seagrass beds are destroyed by coastal construction, land reclamation, shoreline hardening, and dredging (Erftemeijer and Lewis, 2006); damaging fisheries practices such as trawling and aquaculture (Pergent-Martini et al., 2006) also harm seagrass habitats. Mechanical damage from boats, boat moorings, and docks is a problem in some regions (Burdick and Short, 1999; Kenworthy et al., 2002), as are introduced

species (Williams, 2007) that compete for space and resources (Heck et al., 2000). Diseases, such as wasting disease, threaten some seagrasses and have caused large-scale declines (Rasmussen, 1977; Short et al., 1986). Many of the threats are cumulative and some are not mutually exclusive (e.g., most coastal development affects water quality). The effects of global climate change on seagrasses are just beginning to be understood (Short and Neckles, 1999; Waycott et al., 2007; Palacios and Zimmerman, 2007; Björk et al., 2008); and are difficult to document, but whether they manifest as sea level change, heat stress, radiation exposure, or increased storm activity, all largely diminish seagrass habitat, distribution and diversity (Short and Neckles, 1999). However, localized impacts to seagrass species will decrease their survival capacity in the face of global threats.

1.2-Posidonia oceanica

1.2.1-Geographic distribution

Posidonia oceanica (L.) Delile (subphylum Angiospermae, class Monocothyledonae, order Potamogetonales, family Posidoniaceae) is one of five seagrasses present in the Mediterranean Sea, and is the only one to be "endemic" (Short et al., 2007). In the Mediterranean basin *P. oceanica* is distributed along almost all the coasts (Figure 1.4, Pergent-Martini, 2004) covering a surface area of 25.000-50.000 Km^2, representing 1-2% of the total surface (Pasqualini et al., 1998). Recent studies show the presence of *Posidonia oceanica* also in the Marmara Sea (Meinesz et al., 2009) while it is not present along the coasts of Egypt (to the east of the Nile delta, Por, 1978) and in the Black Sea (Gobert et al., 2006). It is rare or absent in the north

Adriatic, in particular along the Italian coasts (Procaccini et al., 2003), and along the coasts of the Camargue and Port-la-Nouvelle (Bouderesque and Meinez, 1982; Boudouresque et al.,2006).

Figure 1.4. *Posidonia oceanica* geographic distribution (highlighted in yellow, modified from Pergent-Martini, 2004).

Throughout its coastal distribution *P. oceanica* forms extensive monospecific meadows (Procaccini et al., 2003, Figure 1.5) that are recorded from 0.5 to 40 m depth (Bouderesque and Meinesz, 1982; Pergent et al., 1995). In turbid waters, the lower limit of the beds is in the order of 9 m (Pergent, 1995).

In the Mediterranean Sea, *P. oceanica* meadows are exposed to a broad range of light and temperature (13-26°C) and show decreasing plants growth and marked decrease in density with depth, ranging from 1.200 shoots m^{-2} in shallow stands to fewer than 100 shoots m^{-2} at depths around 30 m. Water temperature and light vary seasonally inducing seasonal dynamics in shoot biometry and consequently on the overall biomass of the bed (Gobert et al., 2006).

Figure 1.5. *Posidona oceanica* meadow.

1.2.2-P. oceanica architecture

P. oceanica, as well as all higher plants, is provided with differentiated vegetative structures: roots, stems or rhizome and leaves (Figure 1.6). The plant can reproduce both by vegetative growth and sexual reproduction. By "stolonization" (vegetative growth), rhizomes can grow both horizontally (plagiotropic rhizome) and vertically (orthotropic rhizome) (Mazzella et al., 1986).

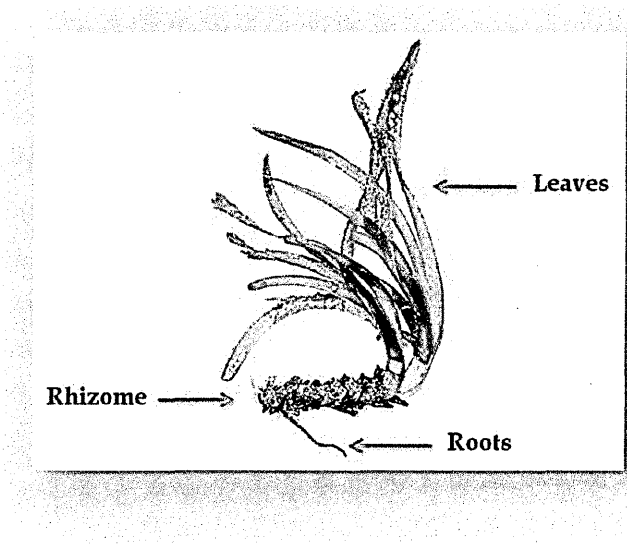


Figure 1.6. *P. oceanica* plant, with leaves, roots and rhizome.

Plagiotropic rhizomes, that allow the expansion-growth of the plant in the horizontal direction, on the lower side have roots thus fulfill both the function of anchoring the plant to the substrate, both the transport of nutrients. The horizontal growth is prevalent because it allows the progressive occupation of the substrate. In certain environmental conditions, for example when the leaf density becomes high and the substrate availability decreases, the rhizomes can activate a growth in the vertical direction allowing the plant to exploit the light and counteract the progressive silting due to the continuous sedimentation of detritus (Boudouresque et al., 1984). An orthotropic rhizome can become plagiotropic and vice versa depending on the space available (Caye, 1980). The horizontal and vertical growth of rhizomes, and the slow decay of this material, causes *Posidonia oceanica* to form a biogenic structure called "matte" (Figure 1.7), that arises from the bottom up to a few meters and can be thousands of years old (Peres and Picard, 1964).

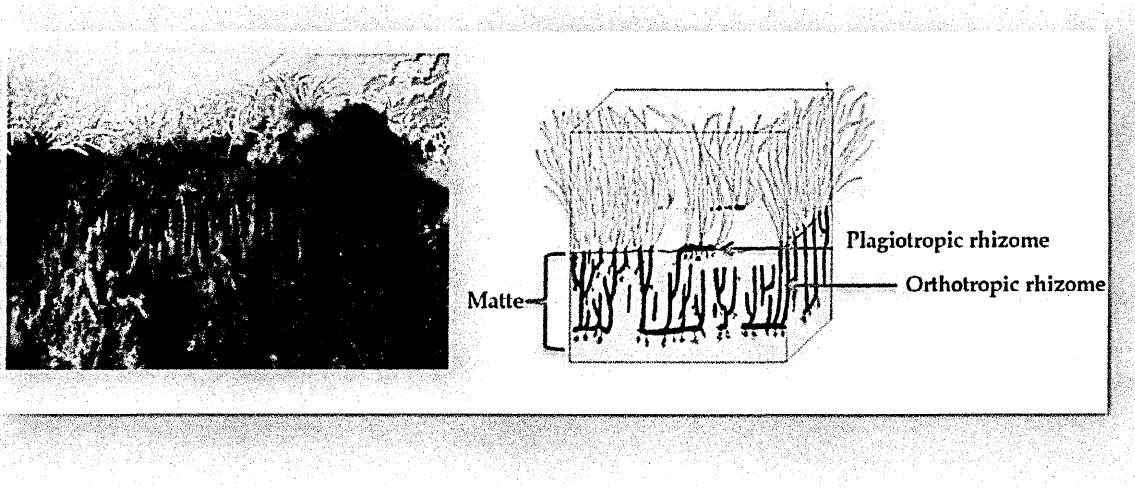


Figure 1.7. *P. oceanica* "matte". Rhizomes can grow both horizontally (plagiotropic rhizome) and vertically (orthotropic rhizome).

The leaves of *Posidonia oceanica* are ribbon-like, with rounded apices (Figure 1.6), have an average width of 1 cm and can reach over 1 m in length. They function both in the photosynthetic process both in the absorption of nutrients from the water (Fresi and Saggiomo, 1981). The leaves are divided into mature, intermediate and young according to their length and the presence of a lignified base (Giraud, 1977). Leaves are gathered in bundles, each of which contains an average of 6-7 leaves: starting from the center of the clump and progressing towards the outside, the leaves increased from young to intermediate up to adults (Panayotidis and Giraud, 1981). The boundle renews itself from the inside out, with a peak in the spring. The growth rate decreases during the summer, when the leaves are longer, covered with epiphytes and begin to manifest a more advanced rate of senescence, characterized by presence of tissue of brown color, photosynthetically inactive (Buia et al., 1992). The leaves, at the height of

senescence, fall mainly in autumn (Romero et al., 1992; Chessa et al., 2000). The detached leaves may decompose within the meadow but also to be transported, through the currents, to other ecosystems or piled on the shoreline. In some areas, particularly on the sandy beaches with gently sloping, accumulations of leaves called banquettes (Figure 1.8, Boudouresque and Meinesz, 1982) may be overstated and arrive at a thickness of 2 m (Boudouresque et al., 2006; Guala et al., 2006 ; Mateo et al., 2003).

Figure 1.8. *Posidonia oceanica* banquette.

1.2.3-Reproductive strategies

Posidonia oceanica, as well as all marine angiosperms, reproduce both by vegetative growth and sexual reproduction.

The vegetative growth, referred to "stolonization", is the main propagation strategy. This process is carried out with both the detachment of terminal rhizomes from the parental rhizome, due to necrosis or hydrodynamism (Molinier and Picard, 1952), and by division and elongation of rhizomes. The vegetative reproduction ensures the propagation of the plant so that it can colonize new environments.

Posidonia oceanica is also able to reproduce asexually through the formation of seedlings directly from the inflorescences (Figure 1.9). The pseudovivipary was observed for the first time near the coast of the island of Formentera (Balearic Islands, western Mediterranean) in May 2004 (Ballesteros et al, 2005).

Figure 1.9. Pseudovivipary in the seagrass *Posidonia oceanica*. *In situ* aspect of pseudoviviparous plantlets still attached to the parental shoot (from Ballesteros et al, 2005).

Although, currently, how frequently pseudovivipary occurs in *Posidonia oceanica*, and which are the mechanisms triggering it are unknown, the fact that it happens occasionally suggests that it may contribute significantly to short-distance dispersal and persistence in a species of great ecological importance, whose meadows can persist for thousands of years and whose colonization of new space occurs very slowly (Ballesteros et al., 2005).

Posidonia oceanica also shows a sexual cycle, with formation of hermaphrodite flowers (clustered from 4 to 10 in an inflorescence of green color) and production of floating fruits also called sea olives (Figure 1.10).

Figure 1.10. Flower and fruits of *Posidonia oceanica*/www.corfubenitses.gr

The ripening of fruits can occur in different times according to their insertion on the axis of the inflorescence (Caye e Meisnez, 1984). A temporal mismatch about the time of anthesis was observed also in relation to the plant distribution along the bathymetric gradient (Mazzella and Buia, 1991): there is a shift of about two months between the beginning of the reproductive cycle in the shallow plants (within 15 m depth) compared to the deep ones (over 15 m). In shallow meadows,

flowers appear in September and October and in late autumn is already possible to observe the early stages of fruit development that reach their maturation in the months of March-April. In deep meadows this same cycle is delayed by about two months (Mazzella and Buia, 1991). The fruits, once matured, detach from the mother plant, and, are transported by currents even in places far away from the meadow of origin, to the benefit of dispersal ability of the species. The fruits, if do not end up on the beach, release the seeds, which fall to the bottom and in the presence of favorable conditions germinate directly.

However, although experiments in the laboratory and in the field have demonstrated a high viability of seeds (Balestri et al., 1998a; Balestri et al., 1998b; Balestri and Bertini, 2005; Piazzi et al, 2000), few seeds are able to germinate under natural conditions (Balestri et al., 1998a; Piazzi et al., 1999); many degenerate assuming a brown-blackish color, and persist in this state for several months.

The flowering and fruiting of *Posidonia* have always been considered to be relatively rare, especially in the more northern portions of the Mediterranean (Molinier and Picard, 1952; Buia and Mazzella, 1991; Piazzi et al., 1999). The reproductive success of the species through sexual reproduction seems to be seriously compromised not only for the big number of seeds that undergo degeneration but also by the action of herbivores that damage a large part of the inflorescence (Balestri and Cinelli, 2003). The sporadic sexual reproduction and the low genetic diversity found in some area of the Mediterranean basin (Procaccini et al., 1996; Procaccini and Mazzella, 1998; Arnaud-Haond et al., 2007; Serra et al., 2010) suggest that clonal propagation can be the main reproductive mode for maintaining and expanding established meadows.

With the presence of a mixed reproductive strategy (clonal and sexual) *P. oceanica* takes advantage simultaneously of fruit and seed production (sexual reproduction) and enhanced production of new shoots (clonal reproduction). The two reproductive modes differ in the dispersal distance and the success of establishment. Compared with seed offspring, clonal ramets can only disperse within a much shorter distance and maintain the parent genotype. Contrary to clonal growth, sexual offspring have long-range dispersal, high rate of mortality and low success of establishment. But more importantly sexual reproduction can maintain higher level of genetic diversity. While with the clonal reproduction it is possible to propagate plants with the same genotypes that can be better adapted to a particular environmental condition, only the presence of genetic diversity ensure the survival of plants in a changing environment.

1.2.4-Ecological functions of P. oceanica meadows

Posidonia oceanica meadows are key ecosystems for the functioning and provision of services to the coastal zone and in the Mediterranean basin (Figure 1.11). *P. oceanica* meadows are the ecosystems with greater biomass and productivity of the Mediterranean (Buia et al., 1992; Boudouresque et al., 1984; Ott, 1980, Wittman, 1984). Both below-ground and above-ground biomass values of *P. oceanica* exceed those of other seagrasses, including the Australian *Posidonia* species (Mateo and Romero, 1997). A striking feature is the distinct partitioning of the biomass, mainly directed into the lignified rhizomes, which can account for up to 90 percent of the total (Pirc, 1983; Mateo and Romero, 1997) and production where leaves account

for more than 90 percent (Wittmann, 1984). In an extensive study, net primary production was estimated to range from 130-1284 g dry weight/m^2/year. However, this production is only minimally used for direct consumption by herbivores (Cebrìan et al., 1996). The *P. oceanica* matte not only represents a net sink of carbon and other elements but also, when growing near the surface, can attenuate the wave action. Under such conditions, it has been estimated that the removal of 1 m^3 of matte can cause 20 m of coastal regression (Jeudy de Grissac, 1984). The banquette has an important role in attenuation of waves and in the protection of beaches from erosion (Boudouresque and Meinesz, 1982). In addition, *P. oceanica* meadows are important for oxygenating water, thanks to their photosynthetic activity. It is estimated that 1 m^2 of meadow gives off up to 14 liters of oxygen a day (Bay, 1978) and that, thanks to the exchanges that occur at the interface between water and air, *P. oceanica* can be regarded as a producer of oxygen also for terrestrial environments, in at least some periods year (Frankignoulle et al., 1984). The very high biodiversity found in *P. oceanica* beds is mostly due to the primary role of this seagrass as a multidimensional habitat for organisms directly participating in the system's trophic dynamics (Mazzella et al., 1992). Also, *P. oceanica* meadows are nursery grounds for the juveniles of many commercially important species of fishes and invertebrates. Among the rare or endangered associated species are the endemic sea star *Asterina pancerii*, the sea horse *Hippocampus hippocampus* and the bivalve *Pinna nobilis*: these species are protected, in both Italy and France, or are included among species requiring a specific legislation for protection (Boudouresque, 1991).

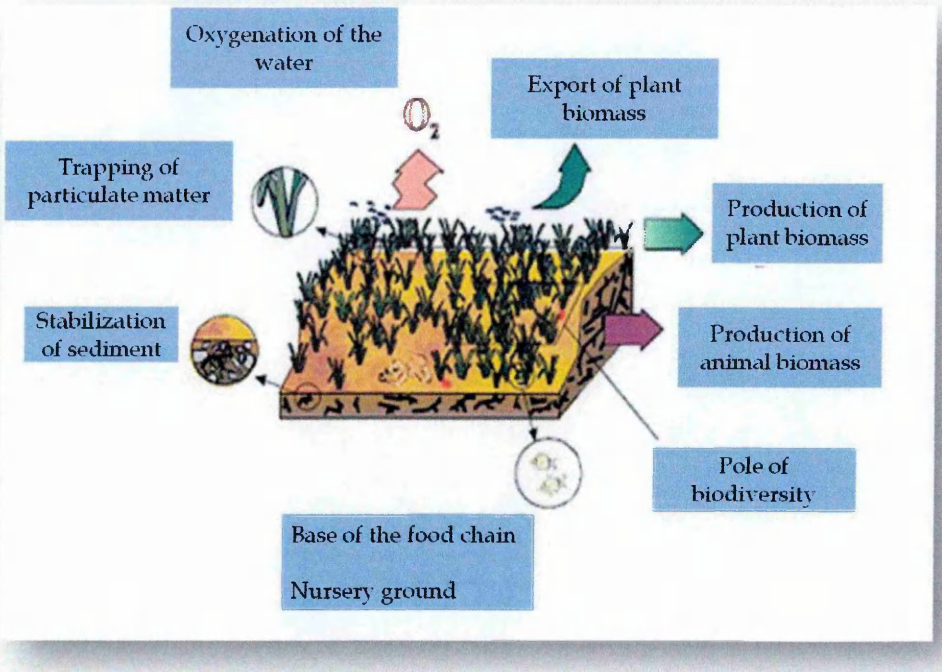


Figure 1.11. *P. oceanica* ecological roles.

1.2.5-Threats

Since the 1980s, 102 of a total of 176 *P. oceanica* meadows reported in the Mediterranean basin have suffered a decline in their covering and/or in the abundance of shoots. The rate of decline of *P. oceanica* meadows is currently 5% per year (Marbá, 2009) due to multiple pressures principally represented by disturbance of anthropogenic origin (eutrophication, disturbance of sedimentary dynamics, the mechanical destruction of the coastal areas, aquaculture, illegal fishing, tourist exploitation, yachting, excessive urbanization etc.) and by fragmentation of natural habitats.

Eutrophication of the coastal area is the main cause of seagrass meadow decline. Dumping nutrients and organic matter into the sea degrades water quality,

principally reducing the amount of light reaching the seagrass meadows and consequently affecting seagrass survival. In particular, seagrasses are extremely vulnerable to any deterioration of water transparency because they exbith the greatest light requirement of all marine photosynthetic organisms (Duarte, 1995). Eutrophication also causes excessive sediment enrichment by nutrient and organic matter. *P. oceanica* meadows are very sensitive to deteriorating sediment quality, and their decline accelerated when the amount of organic matter and phosphorous reaching the sediment exceed 1-2 g dry weight m^{-2} d^{-1} and 0.04 g P m^{-2} d^{-1}, respectively (Dìaz-Almela et al., 2008). *P. oceanica* meadows are acutely sensitive to hydrogen sulfide (Calleja et al., 2007). When the concentration of hydrogen sulfide in the sediment is high and that of oxygen in the plant is low, the hydrogen sulfide in the interstitial water penetrates the seagrass tissues and damages their meristems, thereby reducing their growth and survival (Borum et al., 2005).

Predictions of human population growth suggest that by 2025, 50% of the coast will be built-up area, Mediterranean coastal cities of over 100.000 inhabitants will be home to 90 million residents, and the coastal areas will receive more than 300 million tourists (Benoit and Comeau, 2005). Not only that, aquaculture production in the Mediterranean is expected to double in the next 25 years, maritime freight traffic to almost quadruple, and passenger traffic to likewise double (Benoit and Comeau, 2005). Assuming this scale of growth of human activity, the risk of deterioration of the coastal areas, as exemplified by problems of eutrophication, erosion, proliferation of invasive species, and seabed destruction, and therefore, the risk of seagrass meadow decline, can only increase in the years to come.

In addition, there is also growing evidence that *P. oceanica* meadows are vulnerable to climate change, particularly through the impacts of elevated seawater temperature on shoot survival. Marbá and Duarte (2009) observed a steep *P. oceanica* shoot mortality when seawater exceeds 28°C associated with the occurrence of heat waves in the Mediterranean, which were observed twice along 2002–2007. Whereas the shallow margins of the meadow were able to recover their shoot abundance one year after the heat wave event, net shoot losses persisted for the following years in deeper areas.

In this context of critical environmental conditions has become of crucial importance to develop approaches that could enable to forecast species resistance and resilience, and to help in developing adequate management and conservation strategies. Genetic diversity plays an important role in this scenario. Genetic diversity, in fact, in addition having effects on primary production, inter-specific competition, community structure and fluxes of energy and nutrients also influences population stability under disturbances by reducing realized disturbance, by increasing resistance (decreased biomass loss) or resilience (recovery to pre-disturbance conditions, Hughes et al., 2008; Hughes and Stachowicz, 2010). In the eelgrass *Zostera marina*, Reusch et al. (2005) provided evidence that genotypic diversity enhances ecosystem recovery following an unprecedented heat wave that hit Europe in 2003 and occurred during the manipulative field experiments. Recovery following the heat wave was much better in plots with multiple genotypes (policultures) that in plots with only one genotype (monoculture). The number of shoots was 34% higher in the six-genotypes plots than the monocultures, and the biomass was 26% greater. In

addition, there were benefits in terms of higher abundance (but not diversity) for several eipfaunal groups associated with seagrass. Genotypic diversity also enhanced the resistence of the *Zostera marina* ecosystem to the grazing disturbance where the number of shoots remaining after grazing by geese rose with the increasing genetic diversity (Hughes and Stachowicz, 2004)

The knowledge of the genetic diversity of *P. oceanica* populations can be useful not only to establish the current state of the resource but also to program events of recovery and transplants.

1.2.6-Genetic diversity of P. oceanica in the Mediterranean basin

Posidonia oceanica populations, over the years, have been the subject of several genetic studies that lead to different results depending on the molecular markers used. The first molecular studies about *P. oceanica* using allozymes (Capiomont et al., 1996), M13 DNA fingerprint (Procaccini et al., 1996) and random amplification of polymorphic DNA (RAPD, Procaccini and Mazzella, 1996) indicated an almost complete uniclonality of the meadows, a low number of alleles per locus and an extreme genetic differentiation among populations. Following, the use of new primers RAPD (Jover et al., 2003) indicated a higher level of clonal diversity, contradicting the results previously obtained. This is not surprising for allozymes (i.e. electrophoretically separable morphs of enzymes) considering that they cannot take into account silent mutations in coding and non-coding regions of DNA and, therefore, they underestimate the real genetic variation. With multilocus DNA fingerprinting techniques, instead, polymorphism can be wiped

out from the analysis together with faint or not well represented bands in replicated electrophoretic profiles. The study of genetic structure of *P. oceanica* meadows has seen significant advances with the use of microsatellite markers. The first microsatellite-based surveys of genetic diversity in *P. oceanica* date back to 1998 when six polymorphic loci (Procaccini and Waycott, 1998) were utilized to address meadow diversity of *P. oceanica* in the West Mediterranean basin (Procaccini and Mazzella, 1998). Most of the loci were characterized by trinucleotide repeats (five loci were nuclear and one was from the chloroplast). The same loci were later used in broader scale surveys within the Mediterranean Sea (Procaccini et al., 2001, 2002). These studies detected low levels of variation within *P. oceanica* meadows, with average clonal diversity, calculated over 33 populations and almost 1000 samples, of 0.32 (G/N, where G is the number of genotypes and N is the number of individuals per population). Although the low variation, it was nonetheless evident that populations from the western basin were genetically more polymorphic than those from the eastern Mediterranean. The finding of variation, albeit low, was very interesting because it was significantly higher than the variation detected with the techniques previously applied. However, the low variability remained puzzling when compared with high variability found in the first microsatellite survey of an another seagrass species, *Zostera marina*. For this reason, it was decided to develop new dinucleotide loci (Alberto et al., 2003) which have indeed revealed much higher levels of diversity that are more similar to those found for *Z. marina*. Further comparisons of the two sets of *Posidonia* loci, performed by Serra et al. (2007) have revealed that the markers are not homogeneous and that the combined use of markers, selected on

the basis of their discriminating power, allow to reveal higher genetic diversity in comparisons to the use of single set of markers. A complete reanalysis of *P. oceanica* using all 13 microsatellites across 34 populations and 1200 individuals (Arnaud-Haond et al., 2007) has doubled the earlier estimate of genetic/genotypic diversity, although confirming patterns of genetic subdivision previously found within the Mediterranean basin (Procaccini et al., 2002). In conclusion, if microsatellite diversity is consistently low, then this may rather be an attribute of the markers chosen than of the populations under study, as the two sets of microsatellite markers developed for *P. oceanica* have convincingly shown. Trinucleotide repeats and chloroplast derived loci, in particular, may be less polymorphic than dinucleotide microsatellites.

1.2.7-Genetic structure of Posidonia oceanica at basin and meadow scale

Studies conducted on about 50 *P. oceanica* populations, by using the total set of microsatellite markers, showed that there is a clear separation between western and eastern populations in the Mediterranean (Arnaud-Haond et al., 2007, Serra et al., 2010), confirming the general pattern previously observed with the first set of markers (Procaccini et al., 2002). This analyses indicated also the presence of a third group represented by the populations from Sicily Channel, particularly Pantelleria, Marzamemi and Malta (Arnaud-Haond et al., 2007, Rozenfeld et al., 2008) showing the Sicily Channel as an important genetic boundary between eastern and western Mediterranean basin. The precise position of the transition zone has been assessed more recently, analyzing populations from Calabria coasts

and a bigger number of populations from Sicily Channel (Serra et al., 2010). This analysis shifted the transition zone between the eastern and the western basin, at level of the eastern side of the Sicily Channel (Serra et al., 2010). Thus, these studies highlighted a clear separation between western and eastern populations, originated during last glacial period, when the Strait of Sicily was only a narrow channel (Thiede, 1978). After sea level started to rise (after 18000 years BP), the Strait of Sicily became larger and deeper, and a new connection opened, the very narrow Siculo-Italian Strait (Strait of Messina). Thus, only quite recently the eastern and western populations came in to contact , although sea-currents driven dispersal seem to ensure only limited gene flow between the two sides of the basin (Serra et al., 2010).

At meadow-scale, the analysis of spatial genetic structure was performed in a large *P. oceanica* meadow located in Lacco Ameno (Ischia, Gulf of Naples) that extends along a wide depth gradient (Migliaccio et al., 2005). Results showed that genetic diversity was not evenly distributed throughout the meadow, with consistently higher genetic diversity at the deep margins of the meadow and at the shallow stations. Higher clonality was recorded at stations located in the intermediate positions. Furthermore, different sample groups were identified within the meadow by means of a Bayesian approach. One big group included samples from shallower stations, while all the samples collected in the deep site of the meadow clustered in another group.

The authors thought that the genetic distinction between shallow and deep stations was related to the temporal shift in the ripening period for sexual structures at different depths (Buia and Mazzella 1991), as has already been

suggested by Procaccini et al. (2001). In addition, the high heterozygosity recorded in the meadow could indicate adaptation and should facilitate persistence in the rapidly changing coastal environment. This findings opened new challenges aimed at assessing the extent to which functional plasticity is present and to determine the mechanisms of adaptation to environmental features, such as light attenuation, within the same meadow and along the wide depth gradient.

1.3-References

Abal, E. G. and W. C. Dennison (1996). Seagrass depth range and water quality in Southern Moreton Bay, Queensland, Australia. *Marine and Freshwater Research*, 47: 763-771.

Alberto, F., Correia, L., Arnaud-Haond, S., Billot, C., Duarte, C. M., Serrao, E. (2003). New microsatellite markers for the endemic Mediterranean seagrass *Posidonia oceanica*. *Molecular Ecology Notes*, 3(2) : 253-255.

Arnaud-Haond, S., Migliaccio, M., Diaz-Almela, E., Teixeira, S., Van de Vliet, M. S., Alberto, F., Procaccini, G., et al. (2007). Vicariance patterns in the Mediterranean Sea: east-west cleavage and low dispersal in the endemic seagrass *Posidonia oceanica*. *Journal of Biogeography*, 34(6) : 963-976.

Balestri, E., Piazzi, L., Cinelli F. (1998a). Survival and growth of transplanted and natural seedlings of *Posidonia oceanica* (L.) Delile in a damaged coastal area. *Journal of Experimental and Marine Biology and Ecology*, 228: 209-225.

Balestri, E., Piazzi, L., Cinelli, F. (1998b). In vitro germination and seedling development of *Posidonia oceanica*. *Aquatic Botany*, 60: 83-93.

Balestri, E. and F. Cinelli. (2003). Sexual reproductive success in *Posidonia oceanica*. *Aquatic Botany* 75: 21-32.

Ballesteros, E., Cebrian, E., Garcia-rubies, A., Alcoverro, T., Romero, J., Font, X. (2005). Pseudovivipary, a new form of asexual reproduction in the seagrass *Posidonia oceanica*. *Botanica Marina* 48, 175-177.

Bay, D. (1978). Etude "in situ" de la production primarie d'un herbier de Posidonies (*Posidonia oceanica* (L.) Delile) de la baie de Calvi-Corse. These Fac. Sci., Univ. De Liége, 1-251.

Beck, M. W., K. L. Heck, Jr. , K. Able, D. Childers, D. Eggleston, B.M. Gillanders, B.S. Halpern, C. Hays, K. Hoshino, T. Minello, R. Orth, P. Sheridan, M. Weinstein. (2001). The identification, conservation and management of estuarine and marine nurseries for fish and invertebrates. *Bioscience*, 51: 633-641.

Benoit, G., and A. Comeau. (2005). A sustainable future for the Mediterranean: the Blue Plan's environment and development outlook. London: Earthscan.

Björk, M., Short, F. T., McLeod, E., Beer, S. (2008). Managing Seagrasses for Resilience to Climate Change. IUCN, Gland, Switzerland.

Borum, J., Duarte, C. M., Krause-Jensen, D., Greve T. M. (2004). European seagrasses: an introduction to monitoring and management. The M&MS project. 95 pp.

Borum, J., Pedersen, O., Greve, T. M., Frankovich, T. A., Ziemann, J.C., Fourqurean, J.W., Madden, C.J. (2005). The potential role of plant oxygen and sulphide dynamics in die-off events of the tropical seagrass, *Thalassia testudinum*. *Journal of Ecology*, 93 (1): 148-158.

Boudouresque, C. F. and A. Meinesz. (1982). Découverte de l'herbier de Posidonie. Cah. Parc nation. Port-Cros, Fr., 4: 1-79.

Boudouresque, C. F., Jeudy de grissac, A., Olivier, J. (1984). International Workshop on *Posidonia oceanica* beds. Boudouresque C. F., Jeudy de Grissac A., Olivier J. (eds), *GIS Posidonie publ.*, Fr., 1: 1-454.

Boudouresque, C. F., Avon, M., Gravez, V. (1991). Les espèces à protéger en Méditerranée. GIS *Posidonie*, Marseille . 449 pp .

Boudouresque, C.F., Bernard, G., Bonhomme, P., Charbonnel, E., Diviacco, G., Meinesz, A., Pergent, G., Pergent-Martini, C., Ruitton, S., Tunesi, L. (2006). Préservation et conservation des herbiers à *Posidonia oceanica*. RAMOGE. 202 pp.

Buia, M. C. and L. Mazzella. (1991). Reproductive phenology of the Mediterranean seagrasses *Posidonia oceanica* (L.) Delile *Cymodocea nodosa* (Ucria) Aschers and *Zostera noltii* Hornem. *Aquat Bot*, 40: 343–362.

Buia, M. C., Zupo, V., Mazzella, L. (1992). Primary production and growth dynamics in *Posidonia oceanica*. P.S.Z.N. I: *Marine Ecology*, 13 (1) : 2-16.

Burkholder, J. M., Tomasko, D.A., Touchette, B.W. (2007). Seagrasses and eutrophication. *Journal of Experimental Marine Biology and Ecology*, 350: 46– 72.

Burdick, D. M. and F. T. Short. (1999). The effects of boat docks on eelgrass beds in coastal waters of Massachusetts. *Environmental Management*, 23: 231–240.

Calleja, M., Marbá, N., Duarte, C. M. (2007). The relationship between seagrass (*Posidonia oceanica*) decline and porewater sulfide pools in carbonate sediments. *Estuarine, Coastal and Shelf Science*, 73: 583–588.

Cambridge, M. L. and A. J. McComb (1984). The loss of seagrasses in Cockburn Sound,Western Australia, 1: The time course and magnitude of seagrass decline in relation to industrial development. *Aquatic Botany*, 20: 229–243.

Capiomont, A., Sandmeier, M., Caye, G., Meinesz, A. (1996). Enzyme polymorphism in *Posidonia oceanica*, a seagrass endemic to the Mediterranean. *Aquatic Botany*, 54: 265–277.

Caye G. (1980). Sur la morphogénèse et le cycle végétatif de *Posidonia oceanica* (L.) Delile. Thèse Doct., Univ. Aix-Marseille 2, Fr.: 1-121.

Caye, G. and A. Meisnez. (1984). Observations sur la floraison et la fructification de *Posidonia oceanica* dans la baie de Villefranche et en Corse du Sud. International Workshop on *Posidonia oceanica* beds. Boudouresque C. F., Jeudy de Grissac A., Olivier J. (eds), *GIS Posidonie publ.*, Fr., 1: 193-201.

Cebriàn, J., Duarte, C. M., Marbà, N., Enriquez, S., Gallegos, M., Olesen, B. (1996). Herbivory on *Posidonia oceanica*. Magnitude and variability in the Spanish Mediterranean . *Marine Ecology Progress Series*, 130 : 147-155.

Chessa, L.A., Fustier, V., Fernandez, C., Mura, F., Pais, A., Pergent, G., Serra, S., Vitale, L. (2000). Contribution to the knowledge of 'banquettes' of *Posidonia oceanica* (L.) Delile in Sardinia Island. *Proceeding IV° International Seagrass Biology Workshop, Corsica France, Biologia Marina Mediterranea,* 7: 35–38.

Costanza, R, d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S. , O'Neill, R. V., Paruelo, J., Raskin, R. G., Sutton P., van den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature,* 387: 253–260.

Cox, P.A. (1993). Water-pollinated plants. *Sci Am,* 269(4): 50-56.

Dalla Via, J., Sturmbauer, C., Schoenweger, G., Soetz, E., Mathekowitsch, S., Stifter, M., Rieger, R. (1998). Light gradients and meadow structure in *Posidonia oceanica*: ecomorphological and functional correlates. *Mar Ecol Prog Ser,* 163: 267–278.

de Boer, W. F., (2007). Seagrass-sediment interactions, positive feedbacks and critical thresholds for occurrence. a review. *Hydrobiologia,* 591: 5–24.

den Hartog, C. (1970). The seagrasses of the world. North-Holland Publishing Company, Amsterdam.

Dennison, W. C., Orth, R. J, Moore, K. A., Stevenson, J.C., Carter, V., Kollar, S., Bergstrom, P. W., Batiuk, R. A. (1993). Assessing water quality with submersed aquatic vegetation. *BioScience,* 43(2): 86-94.

Díaz-Almela, E., Marbà, N., Alvarez, E., Santiago, R., Holmer, M., Grau, A., Mirto, S., Danovaro, R, Petrou, A., Argyrou, M., Karakassis, I., Duarte, C. M. (2008). Benthic input rates predict seagrass (*Posidonia oceanica*) fish farm-induced decline. *Mar Pollut Bull.,* 56(7): 1332-42.

Duarte, C. M. and K.Sand-Jensen. (1990). Seagrass colonisation: patch formation and patch growth in *Cymodocea nodosa. Mar Ecol Prog Ser,* 65: 193-200.

Duarte, C. M. (1991). Allometric scaling of seagrass form and productivity. *Mar Ecol Prog Ser,* 77: 289-300.

Duarte, C. M. (1995). Submerged aquatic vegetation in relation to different nutrient regimes. *Ophelia,* 41: 87–112.

Duarte, C. M. and C. L. Chiscano. (1999). Seagrass biomass and production: A reassessment. *Aquatic Botany,* 65: 159–174.

Duarte, C. M., Middelburg, J., Caraco, N. (2005). Major role of marine vegetation on the oceanic carbon cycle. *Biogeosciences,* 2: 1–8.

Erftemeijer, P.L.A. and R. R. R. Lewis. (2006). Environmental impacts of dredging on seagrasses: a review. *Marine Pollution Bulletin,* 52: 1553–1572.

Frankignoulle, M., Bouquegneau, J. M., Ernst, E., Biondo, R., Ringo, M., Bay, D. (1984). Contribution de l'activite de l'herbier de Posidonies au metabolisme global de la baie de Calvi. Premiers resultants. International Workshop on *Posidonia oceanica* beds.

Boudouresque C. F., De Grissac A. J. & Olivier J. (eds), *GIS Posidonie publ.*, Fr., 1984, 1: 23-28.

Fresi, E., and V. Saggiomo. (1981). Phosphorus uptake and transfer in *Posidonia oceanica* (L.) Delile. *Rapp. Comm. Int. Expl. Mer Mèdit.*, 27: 187-188.

Giraud, G. (1977). Contribution à la description et a la phenologie quantitative des herbiers de Posidonia oceanica (L.) Delile. Doctorat de spécialité en oceanologie, Université Aix-Marseille II, Fac. Des Sciences de Lumy, Marseille, 1-50.

Gobert, S., Cambridge, M. L., Velimirov, Pergent, G., Lepoint, G., Bouquegneau, J. M., Dauby, P., Pergent-Martini, C., Walker, D. (2006). Biology of Posidonia. In A. W., Larkum, R. J., Orth, & C. M., Duarte (Eds.), Seagrasses: Biology, Ecology and Conservation (pp. 387-408). Springer.

Green, E. P.and F. T. Short. (2003). World Atlas of Seagrasses. University of California Press.

Guala, I., Simeone, S., Buia, M. C., Flagella, S., Baroli, M., De Falco, G. (2006). *Posidonia oceanica* 'banquette' removal: environmental impact and management implicatios. *Biologia Marina Mediterranea*, 13 (4) : 149 - 153.

Harper, J. L. (1977). Population biology of plants. Academic Press, London.

Heck, K. L., Pennock, J., Valentine, J., Coen, L., Sklenar, S. S. (2000). Effects of nutrient enrichment and large predator removal on seagrass nursery habitats: an experimental assessment. *Limnology and Oceanography*, 45: 1041–1057.

Heck, K. L, Hays, C., Orth, R. J. (2003). A critical evaluation of the nursery role hypothesis for seagrass meadows. *Marine Ecology Progress Series*, 253: 123–136.

Hemminga, M. and C. M. Duarte. (2000). Seagrass Ecology. Cambridge (United Kingdom): Cambridge University Press.

Hughes, A.R. and J.J. Stachowicz. (2004). Genetic diversity enhances the resistance of a seagrass ecosystem to disturbance. *Proceedings of the National Academy of Sciences*, 101:8998-9002.

Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Vellend, M., Underwood, N. (2008). Ecological consequences of genetic diversity. *Ecology Letters*, 11: 609–623.

Hughes, A. R., and J. J., Stachowicz. (2010). Seagrass genotypic diversity increases disturbance response via complementarity and dominance. *Journal of Ecology*, doi:10.1111/j.1365-2745.2010.01767.x

Invers, O., Pérez, M., Romero, J. (1999). Bicarbonate utilization in seagrass photosynthesis: role of carbonic anhydrase in *Posidonia oceanica* (L.) Delile and *Cymodocea nodosa* (Ucria) Ascherson. *J Exp Mar Bio Ecol*, 235: 125-133.

Jeudy de Grissac, A. (1984). Effets des herbiers à *Posidonia oceanica* sur la dynamique marine et la sédimentologie littorale . In : Boudouresque CF, Jeudy de Grissac A, Olivier J (eds) International Workshop Posidonia oceanica Beds. *GIS Posidonie Publication*, 1: 437-443

Jover, M. A, del Castillo-Agudo, L., Garcia-Carrascosa, M., Segura, J. (2003). Random amplified polymorphic DNA assessment of diversity in western Mediterranean populations of the seagrass *Posidonia oceanica*. *Am J Bot*, 90: 364–369.

Kenworthy, W. J., Fonseca, M. S., Whitfield, P.E., Hammerstrom, K. K. (2002). Analysis of seagrass recovery in experimental excavations and propeller-scar disturbances in the Florida Keys National Marine Sanctuary. *Journal of Coastal Research*, 37: 75–85.

Kuo, J. and C. den Hartog.(2006). Seagrasses: Biology, Ecology and Conservation Seagrass Morphology, Anatomy, and Ultrastructure, Dordrecht Netherlands: Springer;.

Les, D. H, Cleland, M. A, Waycott, M. (1997). Phylogenetic Studies in Alismatidae, II: Evolution of Marine Angiosperms (Seagrasses) and Hydrophily. *Systematic Botany*, 22(3): 443-463.

Marbà, N., Duarte, C. M. (1994). Growth response of the seagrass *Cymodocea nodosa* to experimental bunal and erosion. *Mar Ecol Prog Ser*, 107: 307-311.

Marbà, N., Cebrian, J., Enriquez, S., Duarte, C. M. (1996). Growth patterns of Western Mediterranean seagrasses: speciesspecific responses to seasonal forcing. *Mar Ecol Prog Ser*, 133: 203-215.

Marbà, N. (2009). Loss of seagrass meadows from the Spanish coast: results of the Praderas project. In: Fundación BBVA (ed). ISBN: 978-84-96515-84: 61-90.

Marbà, N., and C. M., Duarte. (2009). Mediterranean warming triggers seagrass (*Posidonia oceanica*) shoot mortality. *Global Change Biology*, 16(8) : 2366–2375. doi:10.1111/j.1365-2486.2009.02130.x

Mateo, M. A. and J. Romero. (1997). Detritus dynamics in the seagrass *Posidonia oceanica*: Elements for an ecosystem carbon budget . *Marine Ecology Progress Series*, 151 : 43-53.

Mateo, M. A, Sanchez-Lizaso, J. L., Romero, J. (2003). *Posidonia oceanica* banquettes: a preliminary assessment of the relevance for meadow carbon and nutrients budget. *Estuarine, Coastal and Shelf Science*, 56: 85–90.

Mazzella L., Scipione M. C., Fresi E., Buia M. C., Russo G. F., De Mario R., Lorenti M., Rando A. (1986). Le praterie sommerse del Mediterraneo. La Buona Stampa S.p.a, Ercolano, It, 59.

Mazzella, L. and M. C. Buia. (1991). Temporal evolution and comlexity of Mediterranean seagrass ecosystems. *S.It.E. Proceedings*, 17: 293-295.

Mazzella, L., Buia, M. C, Gambi, M. C, Lorenti, M., Russo, G. F, Scipione, M. B, Zupo, V. (1992). Plant-animal trophic relationships in the *Posidonia oceanica* ecosystem of the Mediterranean Sea : A review. In : John DM (ed) Plant-Animal Interactions in the Marine Benthos. Special Volume No . 46 Systematics Association, Clarendon Press, Oxford . pp 165-187.

Meinesz, A., Cirik, Ş., Akcali, B., Javel, F., Migliaccio, M., Thibaut, T., Yüksek, A., et al. (2009). *Posidonia oceanica* in the Marmara Sea. *Aquatic Botany*, 90(1) : 18–22.

Migliaccio, M., De Martino, F., Silvestre, F., Procaccini, G. (2005). Meadow-scale genetic structure in *Posidonia oceanica. Marine Ecology Progress Series,* 304: 55–65.

Molinier, R. and J. Picard. (1952). Recherches sur les herbiers de Phanérogames marines du littoral Méditerranéen français. *Annales de L'Institut Océanographique,* 27(3) : 157-234.

Orth, R. J, Carruthers, T. J. B, Dennison, W. C, Duarte, C. M., Fourqurean, J. W., Jr, K. L. H., Hughes, A. R, Olyarnik, S., Williams, S. L., Kendrick, G. A., Kenworthy, W. J., Short, F. T., Waycott, M. (2006). A Global Crisis for Seagrass Ecosystems. *BioOne,* 56(12): 987-996.

Ott, J., (1980). Growth and production in *Posidonia oceanica* (L.) Delile. P. S. Z. N. I. *Marine ecology,* Germ., 1, 47-64.

Palacios, S. L. and R. C. Zimmerman. (2007). Response of eelgrass *Zostera marina* to CO2 enrichment: Possible impacts of climate change and potential for remediation of coastal habitats. *Marine Ecology Progress Series,* 344: 1–13.

Panayotidis, P., Makris, P., Catsiki, V. A. (1990). Cycle de bioaccumulation du Cu, Cd, et Cr dans les écailles de *Posidonia oceanica. Rapp. Comm. Int. Mer Medit.,* 32(1) : 13.

Pasqualini, V., Pergent-Martini, C., Clabaut, P., Pergent, G. (1998). Mapping of *Posidonia oceanica* using aerial photographs and side scan sonar: Application of the island of Corsica (France), *Estuarine, Coastal Shelf Science,* 47 (3): 359-367.

Pérès, J. M. and J., Picard. (1964). Noveau manuel de bionomique bentique de la mer Mediterranéenne. *Rec. Trav. St. Mar. Endoume,* 31 (47): 1-137.

Pérez, M., Duarte, C. M., Romero, J., Sand-Jensen, K., Alcoverro, T. (1994). Growth plasticity in *Cymodocea nodosa* stands: the importance of nutrient supply. *Aquat Bot,* 47: 249-264

Pergent, G., Pergent-Martini, C., Boudouresque, C.F. (1995). Utilisation de l'herbiera a *Posidonia oceanica* comme indicateur biologique de la qualité du milieu littoral en Mèditerranée: etat des connaissances. *Mésogée,* 54: 3-27.

Pergent-Martini, C. (2004). Un outil d'évaluation de la qualité écologique du milieu littoral pour la DCE « l'herbier a *Posidonia oceanica* » : État d'avancement de la réflexion du « Groupe d'Athènes ». MEDD, Paris, 8 Décembre 2004.

Pergent-Martini, C., Boudouresque, C.F., Pasqualini, V., Pergent, G. (2006). Impact of fish farming facilities on *Posidonia oceanica* meadows: a review. *Marine Ecology,* 27: 310–319.

Piazzi, L., Acunto, S., Cinelli, F. (1999). In situ survival and development of *Posidonia oceanica* (L.) Delile seedlings. *Aquatic Botany,* 63: 103-112.

Piazzi, L., Balestri, E., Cinelli, F. (2000). Grazing on inflorescences of the seagrass *Posidonia oceanica. Botanica Marina,* 43: 581–584.

Pirc, H. (1983). Below ground biomass of *Posidonia oceanica* (L .) Delile and its importance to the growth dynamics . Proceed Int Symp Aquat Macrophytes, Nijmegen . pp 177-181.

Por, F. D. (1978). Lessepsian migrations. The influx of Red Sea biota into the Mediterranean by way of the Suez canal. Springer Verlag publ.,Berlin: x + 1-228.

Procaccini, G., Alberte, R. S., Mazzella, L., (1996). Genetic structure of the seagrass *Posidonia oceanica* in the Western Mediterranean: Ecological implications. *Marine Ecology Progress Series,* 140: 153-160.

Procaccini G. Mazzella L. (1996). Genetic variability and reproduction in two Mediterranean seagrasses. In: Kuo J, Phillips RC, Walker DI, Kirkman H (eds) Seagrass biology: Proceedings of an International Workshop. Rottnest Island, Western Australia. 25-29 January 1996. University of Western Australia, Nedlands, p 85-92.

Procaccini, G., Mazzella, L. (1998). Population genetic structure and gene flow in the seagrass *Posidonia oceanica* (L.) Delile assessed using a microsatellite analysis. *Marine Ecology Progress Series,* 169: 133-141.

Procaccini, G., Waycott, M. (1998.) Microsatellite loci identified in the seagrass *Posidonia oceanica* (L.) Delile. *Journal of Heredity,* 89: 562–568.

Procaccini, G., Orsini, L., Ruggiero, M. V, & Scardi, M. (2001). Spatial patterns of genetic diversity in *Posidonia oceanica,* an endemic Mediterranean seagrass. *Molecular ecology,* 10(6): 1413–21.

Procaccini, G., M. V. Ruggiero, Orsini, L. (2002). Genetic structure and distribution of microsatellite population genetic diversity in Posidonia oceanica in the entire Mediterranean basin. *Bullettin of Marine Science,* 71(3): 1291-1297.

Procaccini, G., Buia, M., Gambi, M., Perez, M., Pergent-Martini, C., Pergent, G., Romero, J. (2003). The seagrasses of the Western Mediterranean. In: Green EP, Short FT (eds) World atlas of seagrasses. University of California Press, Berkeley, CA, 8: 48–58.

Procaccini, G., Olsen, J. L., Reusch, T. B. H. (2007). Contribution of genetics and genomics to seagrass biology and conservation. *J Exp Mar Biol Ecol,* 350: 234–259.

Rasmussen, E. (1977). The wasting disease of eelgrass (*Zostera marina*) and its effects on environmental factors and fauna. In: McRoy, C. P., Helfferich, C. (Eds.), Seagrass Ecosystems. Marcel Dekker, pp. 1–51.

Reusch, T., Ehlers, A., Hammerli, A., Worm, B. (2005). Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings of the National Academy of Sciences,* 102(8): 2826-2831.

Romero, J., Pergent-Martini, C., Mateo, M. A., Regnier, C. (1992). The detritic compartment in a *Posidonia oceanica* meadow: litter features, decomposition rates and mineral stoks". *Marine Ecology,* 13 (1): 73-83.

Rozenfeld, A. F, Arnaud-Haond, S., Hernandez-García, E., Eguíluz, V. M., Serrão, E. A., Duarte, C. M. (2008). Network analysis identifies weak and strong links in a metapopulation system. Proceedings of the National Academy of Sciences of the United States of America, 105: 18824–18829.

Serra, I. A, Procaccini G., Intrieri, M.C., Migliaccio M., Mazzuca S., Innocenti A.M. (2007). Comparison of ISSR and SSR markers for analysis of genetic diversity in the seagrass *Posidonia oceanica*. *Marine Ecology Progress Series*, 338: 71–79.

Serra, I. A., Innocenti, A.,M., Di Maida, G., Calvo, S., Migliaccio, M., Zambianchi, E., Pizzigalli, C., Arnaud-Haond, S., Duarte, C. M., Serrao, E.A., Procaccini, G. (2010). Genetic structure in the Mediterranean seagrass *Posidonia oceanica*: disentangling past vicariance events from contemporary patterns of gene flow. *Molecular ecology*, 19: 557–568.

Short, F. T., Mathieson, A.C., Nelson, J.I., (1986). Recurrence of the eelgrass wasting disease at the border of New Hampshire and Maine, USA. *Marine Ecology Progress Series*, 29: 89–92.

Short, F. T. and H. A. Neckles. (1999). The effects of global climate change on seagrasses. *Aquatic Botany*, 63: 169–196.

Short, F. T., Carruthers, T.J.B., Dennison, W.C., Waycott, M. (2007). Global seagrass distribution and diversity: a bioregional model. *Journal of Experimental Marine Biology and Ecology*, 350, 3–20.

Short, F. T., Polidoro, B., Livingstone, S.R., Carpenter, K.E., Bandeira, S., Bujang, J.S., Calumpong, H.P., Carruthers, T.J.B., Coles, R.G., Dennison, W.C., Erftemeijer, P.L.A., Fortes, M.D., Freeman, A.S., Jagtap, T.G., Kamal, A.H.M., Kendrick, G.A., Kenworthy, W.J., La Nafie, Y.A., Nasution, I.M., Orth, R.J., Prathep, A., Sanciangco, J.C., van Tussenbroek, B., Vergara, S.G., Waycott, M. and Zieman, J.C. (2011). Extinction risk assessment of the world's seagrass species. *Biological Conservation*, 144: 1961-1971.

Suchanek, T. H.,Williams, S. W.,Ogden, J. C.,Hubbard, D. K., Gill, I. P. (1985). Utilization of shallow-water seagrass detritus by Caribbean deep-sea macrofauna: δ 13C evidence. *Deep Sea Research*, 32: 2201–2214.

Terrados, J. (1997). Is light involved in the vertical growth response of seagrasses when buried by sand? *Mar Ecol Prog Ser*, 152: 295-299.

Terrados, J., Duarte, C. M., Kamp-Nielsen, L., Agawin, N. S. R, Gacia, E., Lacap, D., Fortes, M. D., Borum, J., Lubanski, M., Greve, T. (1999). Are seagrass growth and survival constrained by the reducing conditions of the sediment? *Aquatic Botany*, 65(1-4): 175-197.

Thiede, J. (1978). Glacial Mediterranean. *Nature*, 276, 680–683.

Tomlinson. P. B. (1974). Vegetative morphology and meristem dependence. The foundation of productivity in seagrasses. *Aquaculture*, 4: 107-130.

Touchette, B. W. and J. M. Burkholder. (2000). Overview of the physiological ecology of carbon metabolism in seagrasses. *J Exp Mar Bio Ecol*, 250(1-2): 169-205.

Waycott, M., Duarte, C. M., Carruthers, T. J. B., Orth, R. J., Dennison, W. C., Olyarnik, S., Calladine, A., Fourqurean, J. W., Heck, K. L. Jr, Hughes, A. R., Kendrick, G. A., Kenworthy, W. J., Short, F. T., Williams, S. L. (2009). Accelerating loss of seagrasses across the globe threatens coastal ecosystems. Proceedings of the National Academy of Sciences of the United States of America, 106(30): 12377–81.

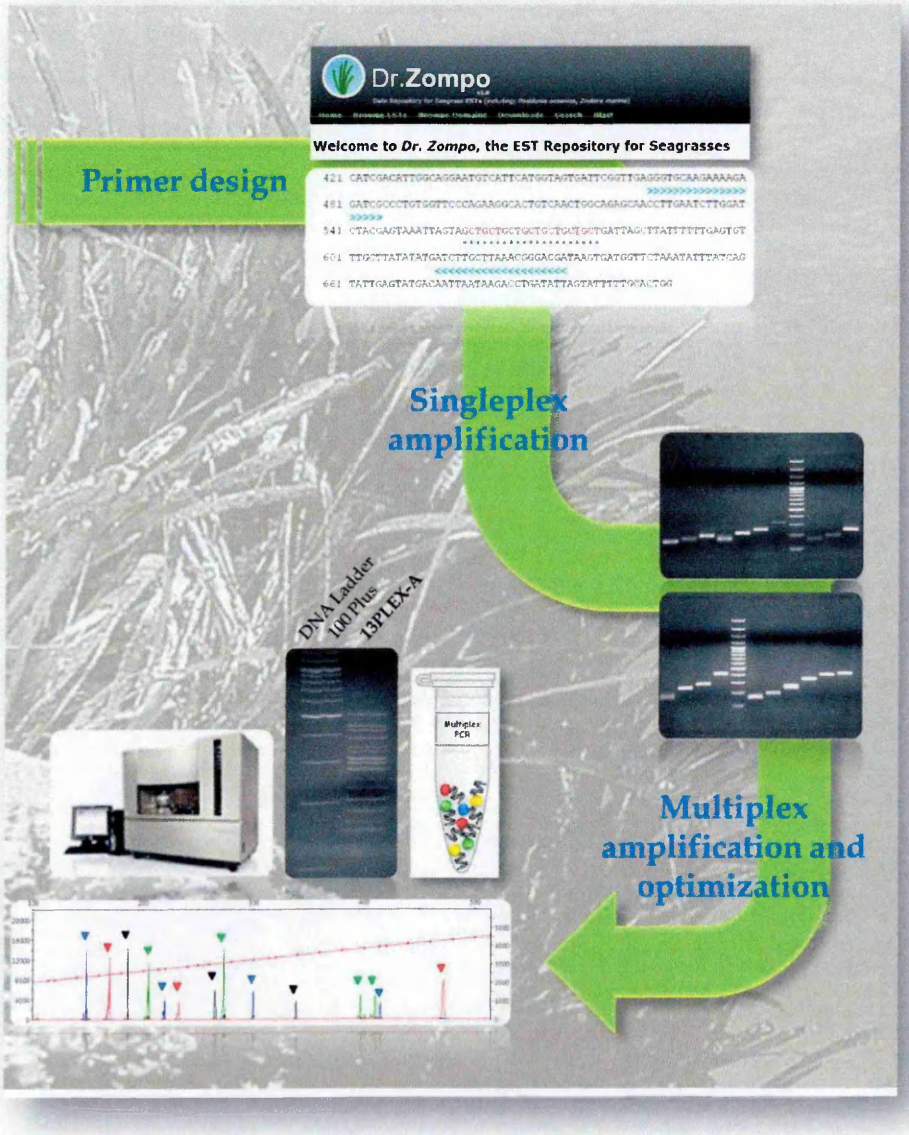
Waycott, M., Collier, C., McMahon, K., Ralph, P., McKenzie, L., Udy, J., Grech, A. (2007). Vulnerability of seagrasses in the Great Barrier Reef to climate change. In: Johnson, J.E., Marshall, P.A. (Eds.), Climate Change and the Great Barrier Reef. Great Barrier Reef Marine Park Authority and Australian Greenhouse Office, pp. 193–236.

Williams, S.L. (2007). Introduced species in seagrass ecosystems: status and concerns. *Journal of Experimental Marine Biology and Ecology* 350, 89–110.

Wissler, L., Codoñer, F. M., Gu, J., Reusch, T. B., Olsen, J. L., Procaccini, G., Bornberg-Bauer, E. (2011). Back to the sea twice: identifying candidate plant genes for molecular evolution to marine life. *BMC Evol Biol.*, 12:11:8.

Wittmann, K. (1984). Temporal and morphological variations of growth in a natural stand of *Posidonia oceanica* (L) Delile *Marine Ecology*, 5 (4) : 301-316.

Chapter 2


Dr.Zompo
Welcome to Dr. Zompo, the EST Repository for Seagrasses
Primer design
Singleplex amplification
DNA Ladder 100 Plus
13PLEX-A
Multiplex PCR
Multiplex amplification and optimization

2-Development of 51 novel EST-linked microsatellites in the Mediterranean seagrass *Posidonia oceanica*

2.1-Abstract

This study reports the isolation, the multiplexing and the characterization of 51 Expressed Sequence Tag (EST)-linked microsatellite markers in the seagrass *Posidonia oceanica*. These markers and 13 anonymous microsatellites previously developed were assembled in 7 multiplex and one singleplex PCR and were tested for their polymorphism in two Mediterranean populations. Of 64 microsatellites analyzed, 26 (14 newly selected within EST regions and 12 anonymous) were polymorphic in at least one of the populations analyzed. The EST-linked markers will be useful for assessing genetic structure and functional diversity in natural populations of *Posidonia oceanica*.

Keywords: local adaptation; conservation; EST-linked microsatellite markers; multiplex PCR; genetic polymorphism; *Posidonia oceanica*.

2.2-Introduction

Microsatellites, or simple sequence repeats (SSRs, Edwards et al., 1991), are tandem repeats of DNA sequences of a few base pairs (1-6 bp) which are widespread in all prokaryotic and eukaryotic genomes analyzed to date (Zane et al., 2002). Several fields of genetic research in plants make use of microsatellites to address various questions such as population genetic analysis, quantitative trait loci (QTL) analysis and traceability (Li et al., 2009). In these contexts, microsatellites are markers of choice, as they have a number of desirable features, such as codominant inheritance, high mutation rates, transferability. In spite of their wide applicability, SSR markers have some drawbacks regarding their lengthy and costly development and their relatively low throughput. However, recently, important progresses have been made in SSR development and genotyping. First, the establishment of genome and expressed sequence tag (EST) sequencing projects in several plant species have generated a vast amount of publicly available sequence data, making the generation of SSRs easier and faster. DNA and EST sequences deposited in on line databases can be easily downloaded and scanned for the identification of SSR, referred as genomic SSRs and EST-SSRs, respectively. Second, multiplexing microsatellites has become much easier by capillary electrophoresis equipment relying on automated laser-induced fluorescence DNA technology (Guichoux et al., 2011).

Recently, Expressed sequence tags (ESTs) derived from two complementary DNA (cDNA) libraries made with samples from a meadow located in Lacco Ameno, Ischia (Gulf of Naples) have been developed for *Posidonia oceanica* and made available in the seagrass database, Dr. Zompo (Wissler et al., 2009). The cDNA

libraries were constructed using pooled total RNA, extracted from 10-15 genotypes from each of the two sampling depths (-5 and -25 m), and were subtracted to select differentially expressed genes. EST libraries has proven to be a useful source of novel, gene-linked molecular markers (Li et al., 2004; Vasemägi et al., 2005) particularly in species in which the full genome has not been sequenced. Microsatellites developed from EST sequence have the advantage of being linked to functional genes, as it is possible to connect the function of the transcript of the genes from an EST sequence with the presence of a microsatellite. This characteristics makes the EST-SSRs particularly useful for genome scans studies for local adaptation. Genome scans are based on the principle of "genetic hitch-hiking"(Maynard Smith and Haigh 1974; Kaplan et al., 1989) according to which if molecular markers are physically linked to functionally important genes, the action of selection on such genes also affects the flanking markers. The genetic hitch-hiking can be easily detected since the selected locus and its linked regions exhibit a variation pattern that is distinct from the rest of the genome.

Here, primers for 51 EST-linked microsatellites were developed and arranged in five multiplexes and one singleplex PCR, reducing genotyping effort for future population genetic and functional genomics studies, such as genome scans.

2.3-Materials and Methods

2.3.1-Collection of samples and DNA isolation

Twenty samples of *Posidonia oceanica* were collected from two localities along the coasts of the Mediterranean basin: Delimara located in the central basin (Strait of Sicily, Malta) and Meloria located in western basin (North Tyrrhenian Sea, Italy, Figure 2.1). In each site, shoots were collected at the reciprocal distance of 7-10 m, minimizing the risk of sampling repeatedly the same genotype (Procaccini et al., 2001). Adult leaf tissues were cleaned from epiphytes and stored at -80°C. Genomic DNA was extracted following the hexadecyltrimethyl ammonium bromide (CTAB) extraction procedure as in Procaccini et al. (1996).

DNA quality was assessed through 1.0% (w/v) agarose/0.5X TBE gels containing 0.5mg/mL ethidium bromide and visualized by UV illumination. DNA purity and concentration was measured using a NanoDrop® ND-1000 Spectrophotometer (Thermo Fisher Scientific).

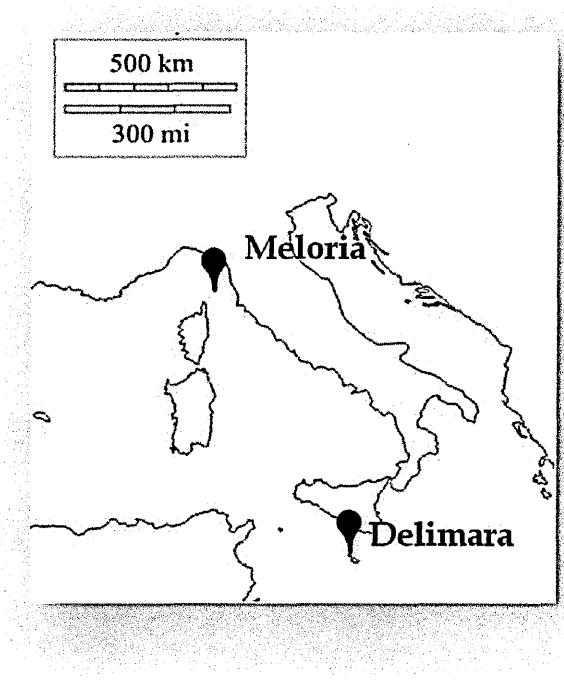


Figure 2.1. Geographical position of the two *Posidonia oceanica* populations sampled. Delimara – Strait of Sicily, Malta (35° 51' 28" N; 14° 33' 07" E); Meloria – Tyrrhenian Sea, Italy (43° 31' 41" N; 10° 10' 32" E).

2.3.2-Microsatellites selection

Two sets of microsatellites, EST-linked and anonymous microsatellites, were used for this study. EST-linked microsatellites were selected considering genes differentially expressed in a *P. oceanica* subtraction cDNA library built using samples collected from a meadow located in Lacco Ameno (Ischia, Gulf of Naples) at two different depths, -5 m (shallow stand) and -25 m (deep stand, Procaccini et al., 2010). A total of 66 contigs containing microsatellites and differently expressed in the two depths were identified, 17 di-nucleotides, 29 tri-nucleotides, 15 tetra-nucleotides, 2 penta-nucleotides and 3 hexa-nucleotides. In addition, when

primers for the selected contigs did not work, contigs containing microsatellites were chosen from Dr. Zompo database as having a long repeat array or involved in photosynthesis and metabolism. Additionally, we used 13 anonymous microsatellites (5 di-nucleotides, 7-tri-nucleotides and one hexa-nucleotide) that had been previously developed (Procaccini and Waycott, 1998; Alberto et al., 2003, highlighted with an asterisk in Table 2.1). They consist of twelve nuclear microsatellites, isolated by screening a *Posidonia oceanica* genomic library, and one chloroplastic obtained from an existing sequence of trnL (UAA) chloroplast intron.

2.3.3-*Multiplex PCR development*

PCR Primer design. To reduce genotyping effort, it was adopted the strategy of microsatellites multiplexing, consisting of the simultaneous amplification of multiple DNA fragments in a single reaction. Thus, the initial phase of PCR primers choice was carefully planned in view of the multiplex PCR design. Compared to singleplex PCR, where only one primer pair is present in the reaction, several primer pairs co-amplify in multiplex PCRs, making primer design and amplification more difficult. First, when planning a multiplex system, primers should be checked during primer design in all combinations to avoid potential formation of primer cross-dimers. Second, primers need to be compatible in terms of annealing temperatures, in order to generate similar quantities of each PCR product. Third, the primers need to generate amplification products within distinct size classes to simplify the scoring of multiple loci using multiplex PCR. Considering these factors, 106 EST-microsatellites were identified as candidate loci

for the study. For this study, perfect, composite and interrupted microsatellites was considered that showed more than 3 repetitions for dinucleotides, trinucleotides and heptanucleotides and more than 2 for tetranucleotides and hexanucleotides. Primer pairs were designed on the sequences available in the seagrass EST database Dr. Zompo, using the web-based software PRIMER3 ver. 0.4.0 (Rozen and Skalestsky, 2000) with the following parameters: optimal primer size of 20 bp (min = 18, max = 25 bp), optimal primer TM of 60°C (min = 57°C, max = 63°C), optimal GC content of 55% (min = 20, max = 80%), product size of 100–550 bp, primer self-complementarity restricted to 8 or fewer bases, and 3' end primer self-complementarity restricted to 3 or fewer bases. In some cases, when the software PRIMER3 did not find suitable primers due to restrictions of the sequence (i.e. when the repeat motif was close to the end of the EST sequence), primers were designed by hand, although presenting parameters slightly different from those above (i.e. lower or higher TM than the range 57°-63°C). .

Simplex Amplification. Each primer pair was initially tested on genomic DNA from one individual in singleplex PCR reactions to check for correct amplification of the desired fragments. After testing different amounts of DNA template, primers concentrations and Taq DNA polymerase units, final singleplex amplifications were performed in a volume of 10 µL consisting of (including) 4 ng of template DNA, 1X PCR Buffer (Roche Applied Science), 5 pmol of each primer, 0.2 mM of each dNTP, and 1.5U of Taq DNA polymerase (Roche Applied Science). For annealing temperature optimization, amplifications were conducted with temperature gradients on a PCR Express Thermal Cycler (Thermo Scientific) with the following reaction profile: initial denaturation for 3 min at 94°C, 35 cycles of

denaturation of 30 s at 94°C, annealing of 40 s at gradient temperature (range: 53°-72°), extension of 1 min at 72°C; and a terminal extension for 10 min at 72°C. All primer pairs were set to work at the annealing temperature of 60°C (Table 2.1). PCR amplification products were analyzed by electrophoresis through 1.5% (w/v) agarose/0.5X TBE gels containing 0.5mg/mL ethidium bromide and visualized by UV illumination. To confirm correct product amplification, PCR products were eluted from the gel with GenElute™ Gel Extraction Kit (Sigma-Aldrich) and sequenced. When primers were designed too close to the repeat motif, sequences were obtained by cloning with TOPO TA Cloning® kit (Invitrogen Corporation; Carlsbad, CA) following the manufacturer's instructions.

Multiplex Amplification and optimization. Microsatellites successfully amplified were selected for the combination in multiplex on the basis of the sizes of amplicons they generated and on the basis of their sequence compatibility. Fast PCR software (Kalendar et al., 2009) was used to analyze the primer pairs compatibility for combination in multiplexes. A specialized multiplex PCR buffer (Qiagen PCR Multiplex kit) was used for multiplex PCR, to increase the specificity of the amplification and to help in overcoming some problems during PCR (i.e. prevents the formation of misprimed products and primer-dimers during reaction setup and ensures comparable efficiencies for annealing and extension of all primers in the reaction), particularly if a high level of multiplexing is targeted. In the first provisional multiplex PCR reactions, all primer pairs were used at equal concentrations and the reaction volumes were the same as suggested by the manufacturer. Multiplex PCR amplifications were conducted in 50 µl reaction volumes containing 1X QIAGEN Multiplex PCR Master Mix (QIAGEN,

Netherlands), 5pmol of each primer pairs (Table 1) and 4 ng of genomic DNA. The recommended Qiagen Multiplex PCR Handbook cycling protocol was used with an annealing temperature of 60°C and 30 PCR cycles. Multiplex thermal cycling was conducted on a GeneAmp® PCR System 9700 (Applied Biosystems) as follows: initial activation of 15 min at 95°C, 30 cycles of denaturation of 30 s at 94°C, annealing at 60°C for 90 s, and extension at 72°C for 60 s. The final extension was at 60°C for 30 min. The inclusion of a final extension step for 30 minutes at 60°C allows the generation of A-overhangs by HotStarTaq DNA polymerase. These overhangs are required for accurate data using capillary DNA sequencers because they minimize the formation of split peaks in the electrophoretic profile. Multiplex PCR amplification products were initially analyzed by electrophoresis through 3.5% (w/v) agarose/0.5X TBE gels and visualized under UV light. The separation of multiplex PCR amplicons on agarose gel allowed to confirm primer pairs compatibility in multiplex design and to have an approximate idea of yields of the single PCR products. After that, a first optimization step started: loci that did not amplify in a particular multiplex were rearranged, primer concentrations were changed, in order to balance PCR products yields and several experiments were performed aimed to reduce the mix volume. Once obtained the desired multiplex amplifications, they were analyzed by capillary electrophoresis for the second optimization step. The separation and detection of PCR products were accomplished with the ABI Prism 3730 automated DNA sequencer (Applied Biosystems, Foster City, CA) using filter dye set G5 (6-FAM™, VIC® NED™, PET®, LIZ® dyes). The internal lane size standard, GeneScan™-600 LIZ®, was labeled with LIZ® dye (Applied Biosystems). Forward primers were labeled with

one of the 4 remaining dyes. Fluorescent dye selection was determined by the amplification-product size of each primer pair, such that there was no size overlapping between amplification products containing the same dye. During the optimization step, multiplex PCR reactions were conducted with the same protocol presented above, except for testing different concentrations of the primer pairs and reduced reaction volumes (see below for the final protocol). After the amplification, multiplexes from the same individuals were analyzed at the automatic sequencer in two conditions: with or without purification of PCR products. About 2 ng of PCR products were mixed with 0.2 µL of internal lane size standard and 12.8 µL formamide (Applied Biosystems) and denatured for 3 min at 95°C. After cooling on ice, the samples were injected in the automated DNA sequencer for 10 s at 2000 V and separated at 15000 V for 44 min at 66°C. Electropherogram profiles were visualized and analyzed using the software PeakScanner (Applied Biosystems). The analysis with the automatic sequencer, that gives quantitative information in the form of peak areas or peak heights, allowed to perform a second, and finer, optimization step regarding the balancing of the amplification yields. Primer concentrations were adjusted stepwise by decreasing those pairs that resulted in relatively stronger signals than the others and increasing the ones producing too weak peaks. The second optimization step resulted in final multiplex PCR conditions (Table 2.1). Multiplex PCR amplifications were conducted in 50 µL (13-Plex-A, 13-Plex-B, 11-Plex and 8-Plex), in 25 µL (5-Plex, 7-Plex and 6-Plex) or in 10 µL (singlePlex) reaction volumes containing 1X QIAGEN Multiplex PCR Master Mix (QIAGEN, Netherlands), equal concentration of each primer pairs (Table 2.1) and 4 ng of genomic DNA. One

locus (Pooc_Contig307), that amplified with very low efficiency, was left out of the multiplex design. The comparison between the electrophoretic profiles from purified and not purified PCR multiplex showed no substantial difference in the resolution of the peaks thus, for the successive capillary electrophoresis, no PCR products purification was applied. The basic strategy for the development of the multiplex PCRs is outlined in Figure 2.2.

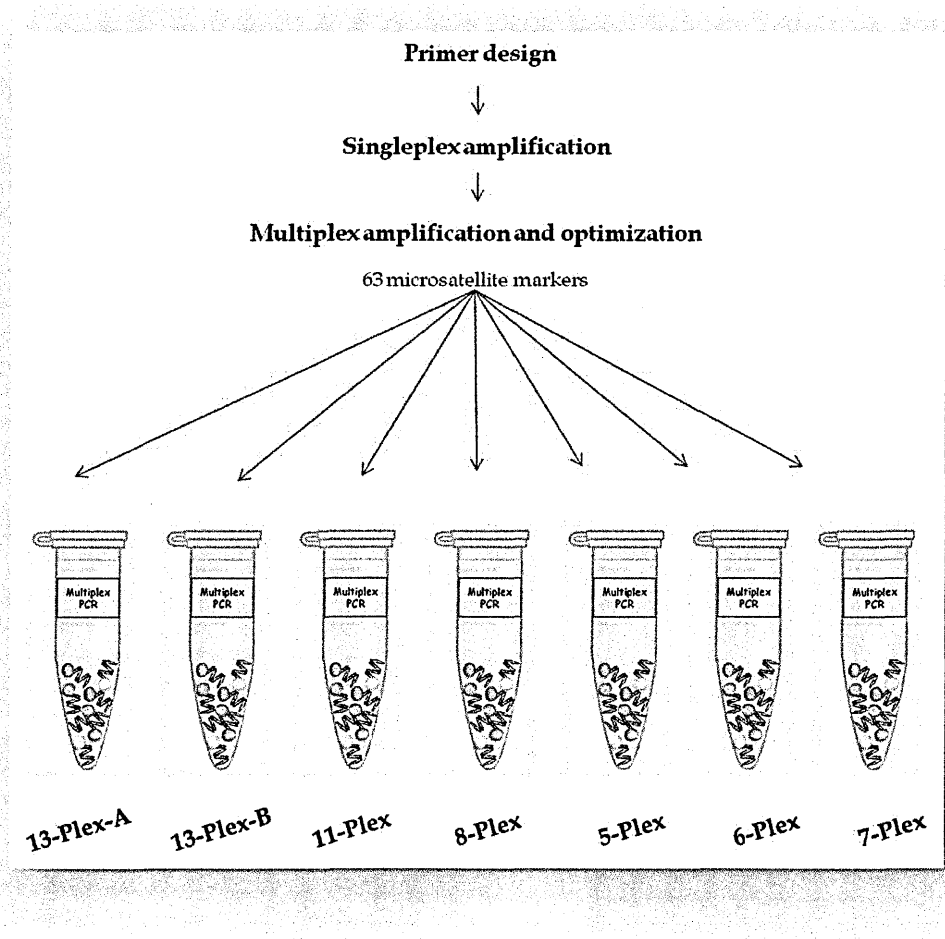


Figure 2.2. Overview of multiplex PCRs development. The 63 microsatellite markers are multiplexed into 7 PCR reactions. 50 EST-linked-microsatellites were combined in five multiplexes, of which two included 13 (13-Plex-A and 13-Plex-B), one 11, one 8, and one 5 primer pairs. One locus (Pooc_Contig307), that amplified with very low efficiency, was left out of the multiplex

design and for this reason is not reported in the figure. 13 anonymous markers were combined in two multiplexes, including 6 (6-Plex) and 7 primers pairs (7-Plex), respectively.

2.3.4-Statistical analysis

The software GenAlEx (Peakall and Smouse, 2006) was used to estimate population genetic parameters: number of alleles (N_a), observed (H_o) and expected (H_e) heterozygosity. Exact test of Hardy-Weinberg equilibrium for each locus in each population was analyzed using the software Arlequin ver. 3.5 (Excoffier and Lischer, 2010). The program GENEPOP v. 3.1 (Raymond and Rousset 1995) was used to calculate linkage disequilibrium between pairs of loci in each population. The *p*-values were adjusted using Bonferroni correction for multiple comparisons with the SGoF program (Carvajal-Rodríguez et al., 2009). Clonal diversity was estimated as:

$$R = \frac{G-1}{N-1}$$

with G representing the number of distinct multilocus genotypes and N representing the total number of samples. Calculations were performed using the software Gimlet (Valière, 2002). MICRO-CHECKER (Van Oosterhout et al., 2004) was employed to test for null alleles and large allele dropout, to score stutter peaks and to estimate null allele frequencies. The neutral locus *Poc-trn* was originally selected from the chloroplastic genome, which is uni-parentally inherited in plants. Hence, it was not utilized for heterozygosity assessment and related statistics.

Table 2.1. Characteristics of 64 microsatellites in *Posidonia oceanica*. 51 were newly selected within EST regions, while 13 were already available. Shown are ID, Repeat Motif (RM), forward and reverse primer name and sequence, Multiplex Reactions (Multiplex), primer amount in pmol, dye. Annealing temperature was always 60°C.

Msat ID	RM	Primer Sequence (5'-3')	Multiplex Reaction	[Primer] pmol	Dye
Pooc-57	$(CAC)_4$	Pooc-57-F: CTTTGGTTGGGGACATTCG Pooc-57-R: GAGACAAGTCATCACTGGTAGTGG	13-Plex-A	4	6-FAM
Pooc-229	$(ACG)_{10}$	Pooc-229-F:CTCAGGGGAAGGTAGAAGAGATG Pooc-229-R:CTGCTTTGTTTCCTCAAAAAGTG	13-Plex-A	4	PET
Pooc-26	$(AAC)_4$	Pooc-26-F: CGACTACCACCACTCGTCATC Pooc-26-R: CTGGGAAGGCAAAGGTAGAG	13-Plex-A	3.2	VIC
Pooc-264	$(GCT)_7$	Pooc-264-F: GCAGGAATGTCATTCATGGT Pooc-264-R: TCGTCCCGTTTAAGCAAGAT	13-Plex-A	12	NED
Pooc-192	$(AT)_7$	Pooc-192-F: TCGGGAAGAAGGAATCCGTCA Pooc-192-R: AGTTAGCACAGGTGCCAC	13-Plex-A	2.4	6-FAM
Pooc-1	$(TCC)_4$	Pooc-1-F: ATCGTCACGTCATCCCCTTT Pooc-1-R: CGCAGAAAGGCGAAAAGGCT	13-Plex-A	4	PET

Pooc-330	$(GAT)_4$	Pooc-330-F: GGAAAACAGCAGCAATCAATC Pooc-330-R: ACGTCTACATCCTCGACCAC	13-Plex-A	5.6	VIC
Pooc-160	$(GA)_5$	Pooc-160-F:CGTCACAGCATACCATTCGC Pooc-160-R:CAACCAATCTCAGATCTCCG	13-Plex-A	4	NED
Pooc-121	$(AAGA)_3$	Pooc-121-F:TCAGATCTTGCGAACCAAAG Pooc-121-R:GAGCAGACACGATAGCCAC	13-Plex-A	3.2	6-FAM
Pooc-PC015C11	$(GAG)_5$ $(CAG)_4$	Pooc-PC015C11- F:ATCACTAGCAGCTGATGGATCTCT Pooc-PC015C11-R:GATCATGTTTTCCTCCCTATG	13-Plex-A	12	VIC
Pooc-145	$(AGA)_4$	Pooc-145-F:AGTCTTCTAGAGGTGCCAATCCA Pooc-145-R:TTCGGTGAGCATTTTATCATACC	13-Plex-A	5.6	PET
Pooc-73	$(ATGC)_3$	Pooc-73-F:CAACTTCCAGAGGGAGAACG Pooc-73-R:TACCCGGCCAATAGTCTGAG	13-Plex-A	3.2	NED
Pooc-3	$(AAG)_7$ $(TAG)_7$	Pooc-3-F:AGCAGAGCTGGGTGAAGAAA Pooc-3-R:CAACATCTCGTCTGTGAATGG	13-Plex-A	5.6	6-FAM
Pooc-322	$(CT)_8$	Pooc-322-F:ACTTCTCCGCGATCTTCTCC Pooc-322-R:CTTGTTTGCACGACACACCT	13-Plex-B	4	6-FAM

Pooc- 321	$(AG)_7$	Pooc-321-F:CACCAGCCTTCTTCTGGATAAAC Pooc-321-R:GAAAAGGTTTTGGGCTTGTAGG	13-Plex-B	4	PET
Pooc-54	$(GGT)_5$	Pooc-54-F:ACATCGTGTTGTTCGCTACCT Pooc-54-R:AAGGCGATCCTCAACTTCC	13-Plex-B	4	NED
Pooc-64	$(AG)_5$	Pooc-64-F:CTCCGAGAATATCTTCTTGTTGGT Pooc-64-R:AACAAGAGCTTTGGAGTCATCAG	13-Plex-B	4	VIC
Pooc-265	$(ATC)_4$	Pooc-265-F: TGTCCTTGTCCTTTGCTTGA Pooc-265-R: CAGCAACCATCTCAAATCCA	13-Plex-B	4	6-FAM
Pooc-153	$(AG)_{16}$	Pooc-153-F:TCCTCCCAATCTAATTTGCTG Pooc-153-R:AAGAACAGTGAGCGCCAGAC	13-Plex-B	4	PET
Pooc-126	$(AAG)_6$	Pooc-126-F:CAATCCAAATCTTTCCCACA Pooc-126-R:CGCCACCATACCGTTCTTC	13-Plex-B	4	NED
Pooc-PC052F01	$(CAG)_4$	Pooc-PC052F01-F:CCCAAGAGACAGGACTGGGTA Pooc-PC052F01-R:CGTGGAGTAGCATCGTAATTGT	13-Plex-B	2.4	VIC
Pooc-96	$(GGT)_4$	Pooc-96-F:CGTCGTCACTGTTCCCAACT Pooc-96-R:CTGTAGAGTTCATAGCTGAA	13-Plex-B	16	6-FAM

Pooc-298	$(CGAACG)_3$	Pooc-298-F:CATGTAACGTGTGCTTGTG Pooc-298-R:GCTCAGAGAGAAAGGCCTGA	13-Plex-B	2.4	NED
Pooc-333	$(TCT)_{10}$	Pooc-333-F:AAAAGTGTTTGAGAGAAACAACTCG Pooc-333-R:ACCTTTGACTTCTCCTTCTGAATCT	13-Plex-B	8	VIC
Pooc-207	$(CT)_7$	Pooc-207-F:GAACTTGCCTCCAACCTCTG Pooc-207-R:ACTGGGGAAAGGTCTGGACT	13-Plex-B	2.4	6-FAM
Pooc-335	$(GGA)_4$	Pooc-335-F:CATGGGCTTCCTCTTGAGAC Pooc-335-R:CTCTCTCATCTCTCGCATCG	13-Plex-B	8	PET
Pooc-PC006A08	$(TG)_5$	Pooc-PC006A08-F:GAAACTCCGATATTGGTCAGCTAT Pooc-PC006A08-R:TAGTCTTTGACATGGCTTAAGTGC	11-Plex	2.4	VIC
Pooc-227	$(CCA)_6$	Pooc-227-F:GCCTCCGACTATGGTAATCTC Pooc-227-R:ACGTTCTCCTTCCGGCTTT	11-Plex	2.4	NED
Pooc-125	$(AG)_{10}$	Pooc-125-F:TTCCTCCCAATCTAATTTGCTG Pooc-125-R:GCTGGTCTGCTGCTTCTGAG	11-Plex	4	6-FAM
Pooc-232	$(TC)_{10}$	Pooc-232-F:ACCACAAAATCAGCTCTCTCTCTC Pooc-232-R:GAGGATTATGACCTTGAGAAGCAT	11-Plex	4	PET

Pooc-PC009A10	$(GGA)_4$	Pooc-PC009A10-F:CACTGCTCCTCCAGTCTTCAT Pooc-PC009A10-R:TGCACAGCACTTCCACTTCCA	11-Plex	2.4	VIC
Pooc-PC045G11	$(GA)_{18}$	Pooc-PC045G11-F:CTCCACCATCCGCCCCCATTTG Pooc-PC045G11-R:CTTCGACAACCTTGCAGGCCCT	11-Plex	4	NED
Pooc-PC046H04	$(TC)_5$	Pooc-PC046H04-F:ACAGTATCCTCTGCAACCCATC Pooc-PC046H04-R:ATCCAGAACCTCCTCCTTGAAC	11-Plex	2.4	6-FAM
Pooc-9	$(TGTT)_3$	Pooc-9-F:GCTGATGTTGATGGTGATGG Pooc-9-R:GCAAATTTCCAAGTCAGAAGAA	11-Plex	4	PET
Pooc-PC025E03	$(CAT)_6$ $(CTT)_5$	Pooc-PC025E03-F:CCATCTGCTGCTGCTTTTCT Pooc-PC025E03-R:GCTCCACGCAATAATCCACT	11-Plex	4	VIC
Pooc-PC044B02	$(CA)_5$	Pooc-PC044B02-F:GCTGCAACTAATACACCTCG Pooc-PC044B02-R:AGGGACGTGATATGCGAGAC	11-Plex	4	NED
Pooc-214	$(CT)_6$	Pooc-214-F:GCCTCCTTCCCCATTGTTCT Pooc-214-R:GCCGAATCCTTCAACTCCAA	11-Plex	4	6-FAM
Pooc-PC020G03	$(TC)_6$	Pooc-PC020G03-F:TTACTTCTCCGCCCTCTC Pooc-PC020G03-R:TTTTCGGCTTCGCGTCTC	8-Plex	1.6	6-FAM

Pooc-300	$(GT)_{12}$	Pooc-300-F:AGAAGGGATGAAGTAAAGATTATG Pooc-300-R:CAACACAAGAGTTGCAAGGG	8-Plex	4	PET
Pooc-61	$(GTT)_4$	Pooc-61-F:CAAGGCACCATTAGGTGGTC Pooc-61-R:AAGGATTTTGCCATGGCTTT	8-Plex	2.4	VIC
Pooc-367	$(TTG)_4$	Pooc-367-F:CGCAAGTTTGAAAGAGCCAG Pooc-367-R:CGGTACTAATGTTATGTCTTC	8-Plex	4	NED
Pooc-PC047G07	$(AGG)_4$	Pooc-PC047G07-F:GCCATCTTGGAAAGCCTAAG Pooc-PC047G07-R:CAGAATCCACTCCCCAACC	8-Plex	2.4	6-FAM
Pooc-16	$(AG)_5$	Pooc-16-F:ACACTTCCTCGTCCACCAGA Pooc-16-R:ACGTCACGATGATAGTACGC	8-Plex	1.6	VIC
Pooc-PC046D10	$(CCG)_4$	Pooc-PC046D10-F:CGTTTCTCGTGCCCAAAG Pooc-PC046D10-R:GTCTTCCCAACATTCCTCGT	8-Plex	8	NED
Pooc-PC031B08	$(CTC)_4$	Pooc-PC031B08-F:TCCATTTTCTCCTCCGTGTC Pooc-PC031B08-R:CACTCTTCGTGCGTCAGAAA	8-Plex	8	PET
Pooc-116	$(AT)_5$	Pooc-116-F:GCTCTTGCCAGAGCATCC Pooc-116-R:AGATGACCAAACACGGCT	5-Plex	8	NED

Pooc-50	$(CAT)_4$	Pooc-50-F:GCTGTGTTTGCCCACTCCT Pooc-50-R:CCTCGATGTGCCAGTTTCTC	5-Plex	4	PET
Pooc-361	$(TC)_{17}$	Pooc-361-F:GTGTTTTGGGCTTGACTAAGAGAT Pooc-361-R:AGGCAACAATTAAGATAGCCCAAAC	5-Plex	4	VIC
Pooc-362	$(GCA)_4$	Pooc-362-F:CCTTTCAATACTCTCTTTTTCCCTA Pooc-362-R:CCTGCTCTTATCCTTTGCCTTAT	5-Plex	2.4	6-FAM
Pooc-PC003H09	$(AG)_8$	Pooc-PC003H09-F:AGCTAGCAAGCTCGTGCA Pooc-PC003H09-R:GACGGTGCTGTTCCACGT	5-Plex	4	NED
Pooc-307	$(AGA)_4$	Pooc-307-F:TCCTTGTAGTCCCTGTCATCG Pooc-307-R:GACATTATTGTCTTTTTGG	1-Plex	16	6-FAM
Poc*-45	$(TCC)_8\,(TTC)_4$	F:AAATTGCCAGATTCTGGTGTCA R:AACCACGTGAGTAAGGGAGGG	7-Plex	4	PET
Po-5**	$(GA)_{20}$	F:CACAATGGCCCGGTAGCA R:GTGGTTGCGTGCCTTCGGTTG	7-Plex	4	VIC
Poc*-5	$(TGG)_{10}$	F:TCTTCGGCTTGTGCTCGTCCTTGA R:GCCTCTCCTGCCCACCACCGCT	7-Plex	2.4	NED

Poc-35*	$(TCC)_{11}$	F:TGGCAAAGTCAATGGCAATAGTAG R:GTACGTCGTCTCGGATGGGAGA	7-Plex	4	6-FAM
Po5-49**	$(GA)_{16}$	F:GGCTCGATGGTGCATTTCAGC R:GCCATTCTCCGCTCTGCTCC	7-Plex	8	PET
Poc-tm*	$(TA)_5$ TTA $(TA)_8$ TAAA $(TA)_4$	F:GGGCAATCCTGAGCCAAATCC R:TTGATATGTCAGTATGTATACGTACG	7-Plex	4	NED
Poc-26*	$(GCGAGGA)_5$	F: GTCACCTTAAATCATCGGAG R:AACGAGGACATAGCGAGTAC	7-Plex	8	PET
Po15**	$(GA)_{20}$	F: AAGCACGCCGCTTAAACCATA R:CATGTTAGTAGGCAATATACTAGGC	6-Plex	4	NED
Po5-10**	$(GA)_{13}$	F:ATGAGACTCCCACAATAACA R:CATGGGAAGGTATAGAAGC	6-Plex	8	PET
Po4-3**	$(GA)_{10}$	F:ACAGAAACTACGAACCATCAG R:TAAGGAGAAGGAGAAGGAAA	6-Plex	8	6-FAM
Po5-39**	$(GA)_{13}$	F:CATTTGGCTGAGTCCCTTTC R:GTCCAAGGCTTCCGTGATGG	6-Plex	1.2	VIC
Po5-40**	$(GA)_{24}$	F:AAAACCAACCCCCACGATAAG R:AATCCAAAGGAACGACACTCA	6-Plex	4	NED

Poc*-42	$(TCC)_8 (TTC)_4$	F: CTCCTTCCTGTACATTCCCCTC R:TGGTCTCTCCCCTTCTCTCACT	6-Plex	0.8	6-FAM

2.4-Results and Discussion

2.4.1-Microsatellite Development

The analytical workflow used for microsatellite marker development resulted in 106 EST-linked microsatellite markers that met the specified criteria and for which primer pairs were designed. From the preliminary screening for amplification success, 51 (48%) microsatellites gave reliable and consistent amplification products, the remaining microsatellites failed to amplify or showed multiple products. The 51 microsatellites include different classes of tandem repeats: 20 di-nucleotides, 27 tri-nucleotides, 3 tetra-nucleotides and 1 hesa-nucleotide (Table 2.1, Materials and Methods). Sixteen of them gave amplicons longer than expected due to the presence of introns, (indicated with asterisks in Table 2.2) as expected for microsatellites designed from EST sequence. Of the 51 microsatellites successfully amplified, 86% were located within ESTs with positive annotation against the SwissProt database (*E*-value from 0.0 to 9.0e-37) while 14% were in ESTs with unknown function (Table 2.2).

Table 2.2. Characteristics of the 51 EST-linked microsatellites selected in *Posidonia oceanica*. For each primer pair, the following parameters are shown: ID, Dr. Zompo (http://drzompo.uni-muenster.de/) Tentative Unigene (TUG) identification with its putative function and Expected value (*E*-value, measure of reliability of sequence similarity) and. GenBank accession numbers. When complete microsatellite sequences are present on more than one EST within a single TUG, all the GenBank Accession. Number. are reported. Sequences containing introns are indicated with asterisks.

Msat ID	TUG ID	Putative Function	*E*-value	GenBank accession number
Pooc-57	Pooc_Contig57	Ethylene-responsive transcription factor 5 (*Arabidopsis thaliana*)	2.0e-26	GO346949, GO348588
Pooc-229	Pooc_Contig229	Thylakoid lumenal protein At1g03610, chloroplastic (*Arabidopsis thaliana*)	5.0e-37	GO347138
Pooc-26	Pooc_Contig26	No Hits	-	GO347513, GO347463, GO348871
Pooc-264	Pooc_Contig264	Nucleoside diphosphate kinase 1 (*Nicotiana tabacum*)	0.0	GO347656, GO347657
Pooc-192	Pooc_Contig192	Photosystem I reaction center subunit IX (*Nymphaea alba*)	3.0e-10	GO346974, GO347699, GO346348
Pooc-1	Pooc_Contig1	Ocs element-binding factor 1 (*Zea mays*)	6.0e-29	GO348601

Pooc-330	Pooc_Contig330	Ferredoxin-1, chloroplastic (*Solanum lycopersicum*)	2.0e-38	GO348399, GO346251
Pooc-160*	Pooc_Contig160	Peroxiredoxin-2B (*Arabidopsis thaliana*)	0.0	JX971045
Pooc-121	Pooc_Contig121	Allene oxide synthase 2 (*Oryza sativa subsp. Japonica*)	0.0	GO346804
Pooc-PC015C11*	Pooc_PC015C11	Protein TIFY 3B (*Arabidopsis thaliana*)	8.0e-11	JX971050
Pooc-145*	Pooc_Contig145	40S ribosomal protein S25 (*Solanum lycopersicum*)	6.0e-26	JX971044
Pooc-73	Pooc_Contig73	Photosystem I reaction center subunit V, chloroplastic (*Arabidopsis thaliana*)	0.0	GO345984
Pooc-3*	Pooc_Contig3	Ribulose bisphosphate carboxylase small chain SSU5B, chloroplastic (*Lemna gibba*)	0.0	JX971041
Pooc-322	Pooc_Contig322	Reticulon-like protein B1 (*Arabidopsis thaliana*)	0.0	GO348835
Pooc- 321	Pooc_Contig321	30S ribosomal protein S20, chloroplastic (*Arabidopsis thaliana*)	4.0e-38	GO347489

Pooc-54	Pooc_Contig54	No Hits	-	GO348703
Pooc-64	Pooc_Contig64	Remorin (*Solanum tuberosum*)	1.0e-29	GO348574, GO348439
Pooc-265	Pooc_Contig265	Metallothionein-like protein type 2 (*Actinidia deliciosa*)	1.0e-13	GO347523, GO347920, GO347792, GO348257, GO348238, GO346390, GO348576, GO348985, GO348420, GO347192
Pooc-153	Pooc_Contig153	Oxygen-evolving enhancer protein 2, chloroplastic (*Fritillaria agrestis*)	0.0	GO347181
Pooc-126*	Pooc_Contig126	No Hits	-	JX971043
Pooc-PC052F01	Pooc_PC052F01	No Hits	-	GO348667
Pooc-96	Pooc_Contig96	Probable signal peptidase complex subunit 1 (*Arabidopsis thaliana*)	5.0e-23	GO346841
Pooc-298*	Pooc_Contig298	Photosystem II reaction center W protein, chloroplastic (*Spinacia oleracea*)	7.0e-07	JX971047
Pooc-333*	Pooc_Contig333	Photosystem II 22 kDa protein, chloroplastic (*Solanum sogarandinum*)	0.0	JX971048

Pooc-207*	Pooc_Contig207	Translation machinery-associated protein 22 (*Chaetomium globosum*)	1.0e-28	JX971046
Pooc-335*	Pooc_Contig335	40S ribosomal protein S15 (*Elaeis oleifera*)	0.0	JX971049
Pooc-PC006A08	Pooc_PC006A08	No Hit	-	GO347794
Pooc-227	Pooc_Contig227	No Hits	-	GO347215
Pooc-125	Pooc_Contig125	Oxygen-evolving enhancer protein 2, chloroplastic (*Fritillaria agrestis*)	0.0	GO347766, GO346855, GO347217
Pooc-232	Pooc_Contig232	Ras-related protein Rab7 (*Gossypium hirsutum*)	0.0	GO346932
Pooc-PC009A10	Pooc_PC009A10	Uncharacterized protein C216.04c (*Schizosaccharomyces pombe*)	2.0e-26	GO348649
Pooc-PC045G11	Pooc_PC045G11	Eukaryotic translation initiation factor 5A-1 (*Nicotiana tabacum*)	0.0	GO348259
Pooc-PC046H04	Pooc_PC046H04	Thioredoxin X, chloroplastic (*Arabidopsis thaliana*)	9.0e-37	GO347795

Pooc-9	Pooc_Contig9	Calmodulin (*Lilium longiflorum*)	0.0	GO346350
Pooc-PC025E03*	Pooc_PC025E03	RING-H2 finger protein ATL4M (*Arabidopsis thaliana*)	0.0	JX971051
Pooc-PC044B02*	Pooc_PC044B02	No Hit	-	JX971051
Pooc-214*	Pooc_Contig214	Multiprotein-bridging factor 1b (*Arabidopsis thaliana*)	0.0	KC012465
Pooc-PC020G03	Pooc_PC020G03	Mitochondrial 2-oxoglutarate/malate carrier protein (*Bos taurus*)	7.0e-38	GO347423
Pooc-300	Pooc_Contig300	Metallothionein-like protein type 3 (*Musa acuminata*)	3.0e-16	GO348869, GO346158, GO348403, GO347380, GO347044, GO347041, GO347037, GO348687, GO347000, GO348213, GO348197, GO348188, GO346089, GO347305, GO347298, GO346412, GO348604, GO345964, GO346746, GO346358, GO348516, GO348511, GO348473, GO348003, GO347941
Pooc-61	Pooc_Contig61	60S ribosomal protein L7-4 (*Arabidopsis thaliana*)	0.0	GO347989

Pooc-367	Pooc_Contig367	60S ribosomal protein L18-2 (*Arabidopsis thaliana*)	0.0	GO348530
Pooc-047G07*	Pooc_PC047G07	50S ribosomal protein L35, chloroplastic (*Spinacia oleracea*)	3.0e-19	JX971055
Pooc-16*	Pooc_Contig16	40S ribosomal protein S15a (*Daucus carota*)	0.0	JX971042
Pooc-PC046D10*	Pooc_PC046D10	40S ribosomal protein S3a (*Cicer arietinum*)	0.0	JX971054
Pooc-PC031B08*	Pooc_PC031B08	Prolyl 4-hydroxylase subunit alpha-1 (*Rattus norvegicus*)	3.0e-22	JX971052
Pooc-116	Pooc_Contig116	Protein vip1 (*Schizosaccharomyces pombe*)	5.0e-13	GO346783
Pooc-50	Pooc_Contig50	14 kDa proline-rich protein DC2.15 (*Daucus carota*)	1.0e-26	GO345974
Pooc-361	Pooc_Contig361	Pathogenesis-related protein 1A (*Nicotiana tabacum*)	2.8e-41	GO347549, GO347125
Pooc-362	Pooc_Contig362	Zeaxanthin epoxidase, chloroplastic (*Solanum lycopersicum*)	4.0e-09	GO348477

Pooc-PC003H09	Pooc_PC003H09	Pathogenesis-related protein PR-1 type (*Sambucus nigra*)	2.0e-11	GO347852
Pooc-307	Pooc_Contig307	Aquaporin PIP1.1 (*Vicia faba*)	0.0	GO348799

Out of 106 markers, 50 EST-linked-microsatellites were combined in five multiplexes, of which two included 13 (13-Plex-A and 13-Plex-B), one 11, one 8, and one 5 primer pairs. One locus (Pooc_Contig307), that amplified with very low efficiency, was left out of the multiplex design. The 13 anonymous microsatellites were arranged in 2 multiplexes including 7 and 6 primer pairs, respectively (Figure 2.3).

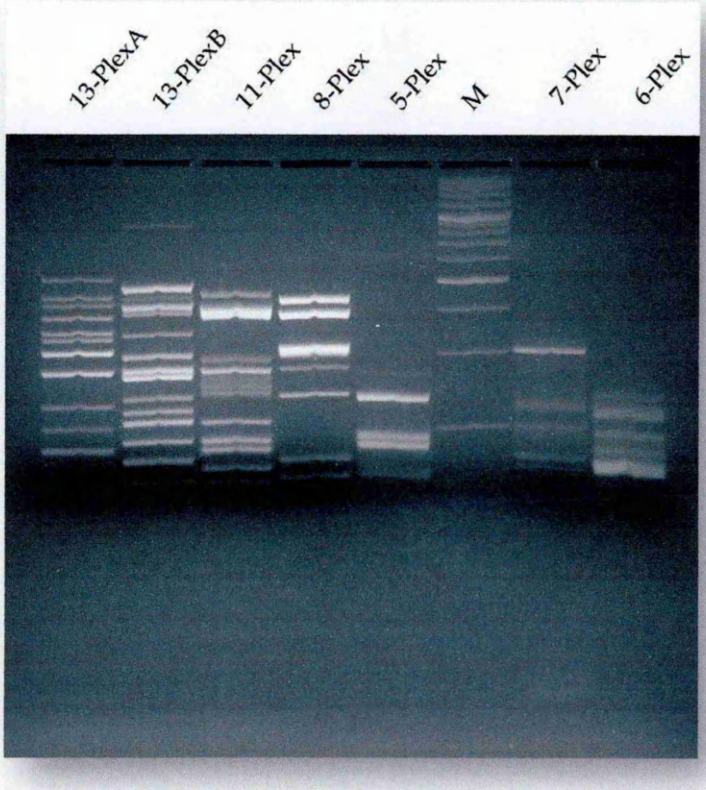


Figure 2.3. Agarose gel electrophoresis of the 7 multiplex PCRs developed. M = DNA ladder 100. Each band represents a locus amplified by PCR. In the lane where the 13-Plex-B was analyzed the presence of an additional PCR product (than 13 PCR products expected) results from the heterozygosity found at locus Pooc-333, where the alleles 397 and 410 were present. Amplification products ranged from 134 bp to 502 bp.

The multiplex optimization steps, consisting mainly of adjustments of primer concentrations, resulted in very clear and balanced electrophoretic profiles (an example of one of the 7 multiplex PCR is showed in Figure 2.4. See appendix, Figure 2.A1 for the remaining 6 multiplex PCR).

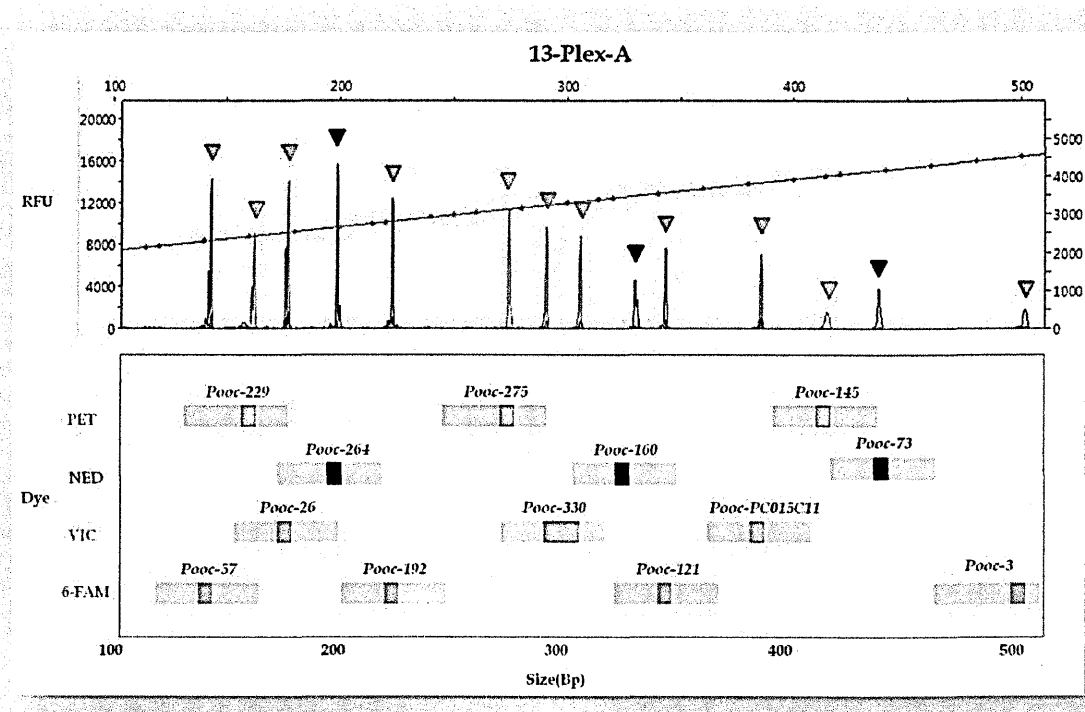


Figure 2.4. Example of a typical electrophoretic profile obtained for the multiplex PCR 13Plex-A for one individual (upper panel) and diagram showing allele size range and fluorescent dyes of each of the multiplexed loci (lower panel). In upper panel, triangles indicate alleles at each locus. In lower panel, dark rectangles represent the observed allele range (base pair); light rectangles represent an arbitrary potential allele size range used during the multiplex design to avoid allele overlap of loci labeled with the same fluorescent dye.

2.4.2-*Microsatellite Characterization*

The intra-population polymorphism of the 64 microsatellites (51 EST-linked and 13 anonymous microsatellites) was tested by genotyping 20 individuals from 2 populations located in Delimara and Meloria. Out of 64 microsatellites analyzed, 26 (14 EST and 12 anonymous) were polymorphic in at least one of the two populations, the remaining 38 were monomorphic (see Figure 2.5 for the 26 polymorphic loci).

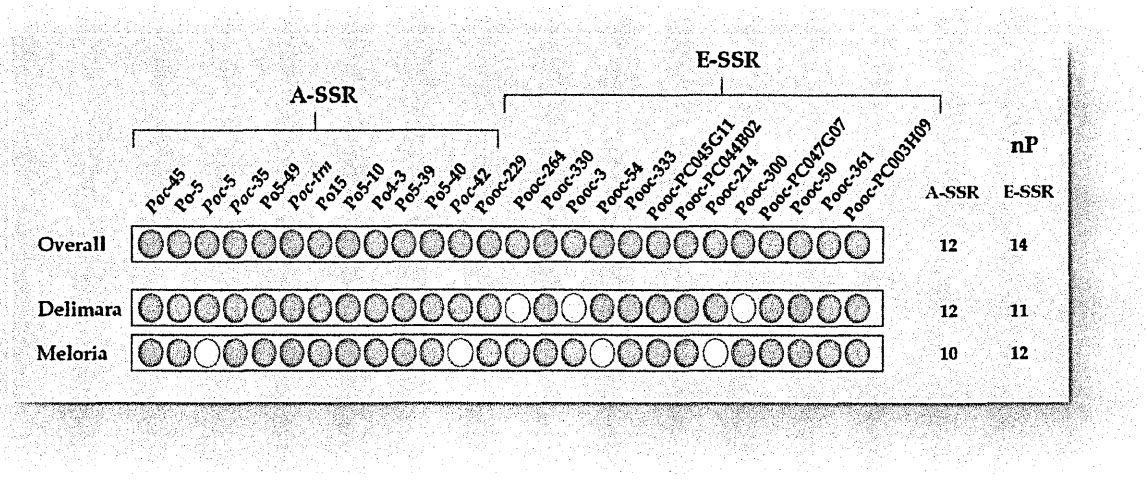


Figure 2.5. Polymorphic loci in the two populations analyzed. Filled circle = polymorphic locus; empty circle = monomorphic locus; nP = number of polymorphic loci; A-SSR= anonymous microsatellites. E-SSR = EST-linked microsatellites.

Twenty-three loci were polymorphic in Delimara, exhibiting from 2 to 10 alleles/locus with an average of 3.391 alleles/locus. Twenty-two microsatellites were polymorphic in Meloria, exhibiting from 2 to 9 alleles/locus with an average of 3.136 alleles/locus (Table 2.3). In Delimara, the mean number of alleles per locus was 3.833 for anonymous markers and 2.909 for E-SSRs. In Meloria, the mean number of alleles per locus was 3.700 for anonymous markers and 2.667 for

E-SSRs. Functional constraints in transcribed versus neutral regions of the genome may result in restriction on the range of potential allelic state at a given locus, and hence lead to a low level of diversity for EST-linked markers. All individuals analyzed showed distinct multilocus genotypes (R = 1 for both populations). Observed (H_o) and expected (H_e) heterozygosities values ranged from 0.150 to 1 (average = 0.647) and from 0.139 to 0.850 (average = 0.472) respectively, in Delimara, and from 0.050 to 1 (average = 0.599) and from 0.049 to 0.723 (average = 0.483) respectively, in Meloria. Significant deviations from Hardy-Weinberg equilibrium expectations were detected for 7 loci in Delimara and for 8 loci in Meloria (indicated with asterisks in Table 2.3). Most of the loci presented an excess of heterozygosity. Heterozygosity excess has been already observed in *Posidonia oceanica* meadows (Migliaccio et al., 2005; Arnaud-Haond et al., 2007; Tomasello et al., 2009; Meinesz et al., 2009) and, in general, seem to be a characteristic of seagrasses: heterozygosity excess has been also found in *Cymodocea nodosa, Zostera noltii* (Ruggiero et al., 2005a,b), and *Zostera marina* (Hämmerli et al., 2003). Within a population, heterozygote excess revealed by negative F_{IS}, has several potential causes. It may result from small reproductive population size. When only few breeders contribute to the next generation, allelic frequencies can differ between male and female parents by chance alone (Rasmussen 1979; Pudovkin et al. 1996). Negative F_{IS} may also arise through negative assortative mating between individuals carrying different alleles. Another explanation for negative F_{IS} involves asexual reproduction (Delmotte et al . 2002; Balloux, 2004; Alberto et al . 2005; Ruggiero et al . 2005a) that maintains heterozygosity or even increases it by mutation over generations (Judson and Normark 1996; Welch and Meselson 2000).

Lastly, high observed heterozygosity could depend from balancing selection leading to local adaptation via heterozygote advantage (overdominance) and thus be correlated with individual fitness (Stoeckel et al., 2006). In theory this explanation should not be valid if the microsatellites used are truly neutral, which is a general assumption for these molecular markers. Nevertheless, it has been showed that microsatellites can have various functions within the genome (Li et al., 2002) hence, it is possible that also some loci used in this study could have patterns not consistent with neutrality.

A general excess of homozygotes for most allele size classes was found at two loci (Po5-40, Pooc-PC045G11) in Meloria, as showed by the analysis with MICRO-CHECKER software. The analysis with MICRO-CHECKER also indicated no scoring errors attributable to stuttering or large allele dropout. Evidence of null alleles was found at two loci, Po5-40 and Pooc-PC045G11, in Meloria. Null allele frequencies for all loci, estimated following Brookfield 1 equation, are reported in Table 2.3. Test for linkage disequilibrium for each pair of loci in each population revealed no significant associations after Bonferroni correction (significance level = 0.05).

Table 2.3. Population genetic parameters for the analyzed populations. The cpDNA locus *Poc-trn* has not been utilized for the assessment of heterozygosities, Hardy-Weinberg equilibrium and for the estimation of null alleles. N = sample size, N_g = number of successfully genotyped individuals, N_a = number of alleles per locus, alleles size range (bp), H_o = observed heterozygosity, H_e = expected heterozygosity, P = P-values of exact test of Hardy-Weinberg equilibrium (asterisks indicate significant departure from Hardy-Weinberg equilibrium, *=P<0.05), NAF = null allele frequency values for all loci following Brookfield 1 estimator, m = monomorphic locus.

	Delimara (N=20)							Meloria (N=20)						
Locus	**N_g**	**N_a**	**Size**	**H_o**	**H_e**	**P**	**NAF**	**N_g**	**N_a**	**Size**	**H_o**	**H_e**	**P**	**NAF**
Pooc-229	20	3	150-168	0.500	0.560	0.19837	0.0385	20	2	162-168	0.550	0.439	0.60723	-0.0773
Pooc-264	20	m	199	m	m	-	0	20	2	190-199	0.050	0.049	1	-0.0012
Pooc-330	20	2	291-306	0.750	0.489	0.06040	-0.1755	20	2	291-306	0.200	0.180	1	-0.0169
Pooc-3	20	m	502	m	m	-	0	20	2	502-505	0.500	0.420	0.61827	-0.0563
Pooc-54	20	2	183-186	0.450	0.349	0.52811	-0.0751	20	m	186	m	m	-	0
Pooc-333	20	3	398-413	0.400	0.335	1	-0.0487	20	3	398-413	0.750	0.585	0.29516	-0.1041
Pooc-PC045G11	20	9	220-250	0.950	0.840	0.00584*	-0.0598	19	9	220-240	0.632	0.834	0.00958*	0.1103

Pooc-PC044B02	20	2	385-387	0.850	0.489	0.00180*	-0.2427	20	2	385-387	0.850	0.499	0.00514*	-0.2344
Pooc-214	20	2	406-408	0.150	0.139	1	-0.0099	20	m	406	m	m	-	0
Pooc-300d	20	m	168	m	m	-	0	20	2	158-168	0.600	0.420	0.11831	-0.1268
Pooc-PC047G07	20	2	290-296	1.000	0.500	0.00001*	-0.3333	20	2	290-296	1.000	0.500	0.00001*	-0.3333
Pooc-50	20	2	178-181	1.000	0.500	0.00001*	-0.3333	20	2	178-181	1.000	0.500	0.00001*	-0.3333
Pooc-361	20	3	178-188	0.250	0.224	1	-0.0215	20	2	186-188	0.300	0.255	1	-0.0359
Pooc-PC003H09	20	2	230-234	0.950	0.499	0.00010*	-0.3011	20	2	230-234	0.900	0.495	0.00076*	-0.2709
***Poc*-45**	20	2	116-140	0.950	0.499	0.00009*	-0.3011	20	2	116-140	1.000	0.500	0.00001*	-0.3333
Po-5	20	2	161-179	0.500	0.420	0.62012	-0.0563	20	3	161-185	0.700	0.549	0.54164	-0.0977
***Poc*-5**	20	2	170-173	0.150	0.139	1	-0.0099	20	m	170	m	m	-	0
***Poc*-35**	20	2	196-199	0.950	0.499	0.00012*	-0.3011	20	3	190-199	0.600	0.445	0.31890	-0.1073
Po5-49	20	7	216-240	0.750	0.735	0.14585	-0.0086	20	5	216-240	0.350	0.533	0.15550	0.1191

Poc-tm	20	2	291-301	-	-	-	-	20	2	301-311	-	-	-	-
Po15	20	6	134-152	0.900	0.750	0.32480	-0.0857	20	6	132-154	0.900	0.763	0.90151	-0.078
Po5-10	20	3	161-165	0.500	0.454	0.38326	-0.0318	20	2	161-163	0.300	0.320	1	0.0152
Po4-3	20	5	160-168	0.450	0.405	0.56679	-0.032	20	4	164-170	0.500	0.656	0.03072*	0.0943
Po5-39	20	3	175-179	0.800	0.579	0.20250	-0.1401	20	3	175-179	0.550	0.486	1	-0.0429
Po5-40	18	10	189-229	0.889	0.850	0.11511	-0.0209	20	7	201-217	0.350	0.723	0.00000*	0.2163
Poc-42	20	2	210-216	0.150	0.139	1	-0.0099	20	M	210	M	M	-	0
Mean		3.391		0.647	0.472				3.136		0.599	0.483		

2.5-Conclusions

Here, the isolation, multiplexing and characterization of EST-linked microsatellites in *P. oceanica* are reported. These markers open new perspectives for functional studies in this species that may lead to the identification of adaptive responses of natural populations to environmental challenges.

2.6-References

Alberto, F., Correia, L., Arnaud-Haond, S., Billot, C., Duarte, C. M., Serrao, E. (2003). New microsatellite markers for the endemic Mediterranean seagrass *Posidonia oceanica. Molecular Ecology Notes,* 3(2): 253–255. doi:10.1046/j.1471-8286.2003.00415.x

Alberto, F., Gouveia, L., Arnaud-Haond, S., Pérez-Lloréns, J. L., Duarte, C. M., Serrão, E.A.. (2005). Withinpopulation spatial genetic structure, neighbourhood size and clonal subrange in the seagrass Cymodocea nodosa. *Molecular Ecology,* 14: 2669–2681.

Arnaud-Haond, S., Migliaccio, M., Diaz-Almela, E., Teixeira, S., Van de Vliet, M. S., Alberto, F., Procaccini, G., Duarte, C. M., Serrão, E. A. (2007). Vicariance patterns in the Mediterranean Sea: east-west cleavage and low dispersal in the endemic seagrass *Posidonia oceanica. Journal of Biogeography,* 34(6): 963–976. doi:10.1111/j.1365-2699.2006.01671.x

Balloux, F. (2004). Heterozygote excess in small populations and the heterozygote-excess effective population size. *Evolution,* 58: 1891–1900.

Carvajal-Rodríguez, A., De Uña-Alvarez, J., Rolán-Alvarez, E. (2009). A new multitest correction (SGoF) that increases its statistical power when increasing the number of tests. *BMC bioinformatics,* 10, 209. doi:10.1186/1471-2105-10-209

Delmotte, F., Leterme, N., Gauthier, J. P., Rispe, C., Simon, J.C. (2002). Genetic architecture of sexual and asexual populations of the aphid Rhopalosiphum padi based on allozyme and microsatellite markers. *Molecular Ecology,* 11: 711–723.

Edwards, A., Civitello, A., Hammond, H. A., Caskey, C. T. (1991). DNA typing and genetic mapping with trimeric and tetramerictandem repeats. *Am J Hum Genet.,* 49: 746–756

Excoffier, L. and H. E. L. Lischer (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources,* 10(3): 564–7. doi:10.1111/j.1755-0998.2010.02847.x

Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., Lepoittevin, C., Malausa, T., Revardel, E., Salin, F., Petit, R. J. (2011). Current trends in microsatellite genotyping. *Molecular ecology resources,* 11(4): 591–611. doi:10.1111/j.1755-0998.2011.03014.x

Hämmerli, A. and T.B.H. Reusch, (2003). Inbreeding depression influences genet size distributionin a marine angiosperm. *Mol. Ecol.* 12: 619–629.

Judson, O.P. and B. B. Normark. (1996). Ancient asexual scandals. *Trends in Ecology & Evolution,* 11: A41–A46.

Kalendar, R., Lee, D., Schulman, A. H. (2009). FastPCR Software for PCR Primer and Probe Design and Repeat Search. *Genes, Genomes and Genomics,* 3: 1-14.

Kaplan, N. L., Hudson, R. R., Langley, C. H. (1989). The "hitchhiking effect" revisited. *Genetics,* 123: 887– 899.

Li, Y.-C., Korol, A. B., Fahima, T., Beiles, A., Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular ecology*, 11(12): 2453–65.

Li, Y.-C., Korol, A. B., Fahima, T., Nevo, E. (2004). Microsatellites within genes: structure, function, and evolution. *Molecular biology and evolution*, 21(6): 991–1007.

Li, M., Barkleyl, N. A., Jenkins, T. M. (2009). Microsatellite Markers in Plants and Insects. Part I: Applications of Biotechnology. *Genes, Genomes and Genomics*, 3; 54-67.

Maynard, S. J. and J. Haigh. (1974). The hitch-hiking effect of a favourablegene. *Genetical Research*, 23: 23– 35.

Meinesz, A., Cirik, Ş., Akcali, B., Javel, F., Migliaccio, M., Thibaut, T., Yüksek, A., Procaccini, G. (2009). *Posidonia oceanica* in the Marmara Sea. *Aquatic Botany*, 90(1): 18–22. doi:10.1016/j.aquabot.2008.04.013

Migliaccio, M., De Martino, F., Silvestre, F., Procaccini, G. (2005). Meadow-scale genetic structure in *Posidonia oceanica*. *Marine Ecology Progress Series*, 304: 55–65.

Peakall, R. and P. E. Smouse. (2006). Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1): 288–295. doi:10.1111/j.1471-8286.2005.01155.x

Procaccinil, G. and S. Randall. (1996). Genetic structure of the seagrass *Posidonia oceanica* in the Western Mediterranean: ecological implications. *Marine Ecology Progress Series*, 140: 153–160.

Procaccini, G. and M. Waycott. (1998.) Microsatellite loci identified in the seagrass *Posidonia oceanica* (L.) Delile. *Journal of Heredity*, 89: 562–568

Procaccini, G., Orsini, L., Ruggiero, M. V, Scardi, M. (2001). Spatial patterns of genetic diversity in *Posidonia oceanica*, an endemic Mediterranean seagrass. *Molecular ecology*, 10(6): 1413–21.

Procaccini, G., Dattolo, E., Innocenti, AM., Mazzuca, S., Cavallini, A., Natali, L. J., Bayer, P. G. (2010). Differential gene expression profiling under different light conditions in *Posidonia oceanica* (L.) Delile by SSH analysis. *Rapp Comm Int Mer Médit.*, 39: 641

Pudovkin, A. I., Zaykin, D. V., Hedgecock, D. (1996). On the potential for estimating the effective number of breeders from heterozygoteexcess in progeny. *Genetics*, 144: 383–387.

Rasmussen, D. I. (1979). Sibling clusters and genotypic frequencies. *American Naturalist*, 113: 948–951.

Raymond, M. and F. Rousset. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86: 248 - 249

Rozen, S. and H. Skalestsky. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology*, (Clifton, N.J.) 132: 365-386.

Ruggiero, M. V., Reusch, T. B. H., Procaccini, G. (2005a). Local genetic structure in a clonal dioecious angiosperm. *Molecular ecology*, 14(4): 957–67.

Ruggiero, M. V., Capone, S., Pirozzi, P., Reusch, T. B. H., Procaccini, G. (2005b). Mating System and Clonal Architecture: A Comparative Study in Two Marine Angiosperms. *Evolutionary Ecology*, 19(5): 487–499.

Stoeckel, S., Grange, J., Fernández-Manjarres, J. F., Bilger, I., Frascaria-Lacoste, N., Mariette, S. (2006). Heterozygote excess in a self-incompatible and partially clonal forest tree species-*Prunus avium* L. Molecular ecology, 15(8): 2109–18.

Tomasello, A., Di Maida, G., Calvo, S., Pirrotta, M., Borra, M., Procaccini, G. (2009). Seagrass meadows at the extreme of environmental tolerance: the case of *Posidonia oceanica* in a semi-enclosed coastal lagoon. *Marine Ecology*, 30(3): 288–300.

Valière, N. (2002). GIMLET: a computer program for analysing genetic individual identification data, *Molecular Ecology Notes*, 4(3): 377–379.

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., Shipley, P. (2004). Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. *Molecular Ecology Notes*, 4(3): 535–538.

Vasemägi, A., Nilsson, J., Primmer, C. R. (2005). Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in atlantic salmon (*Salmo salar L.*). *Molecular biology and evolution*, 22(4): 1067–76.

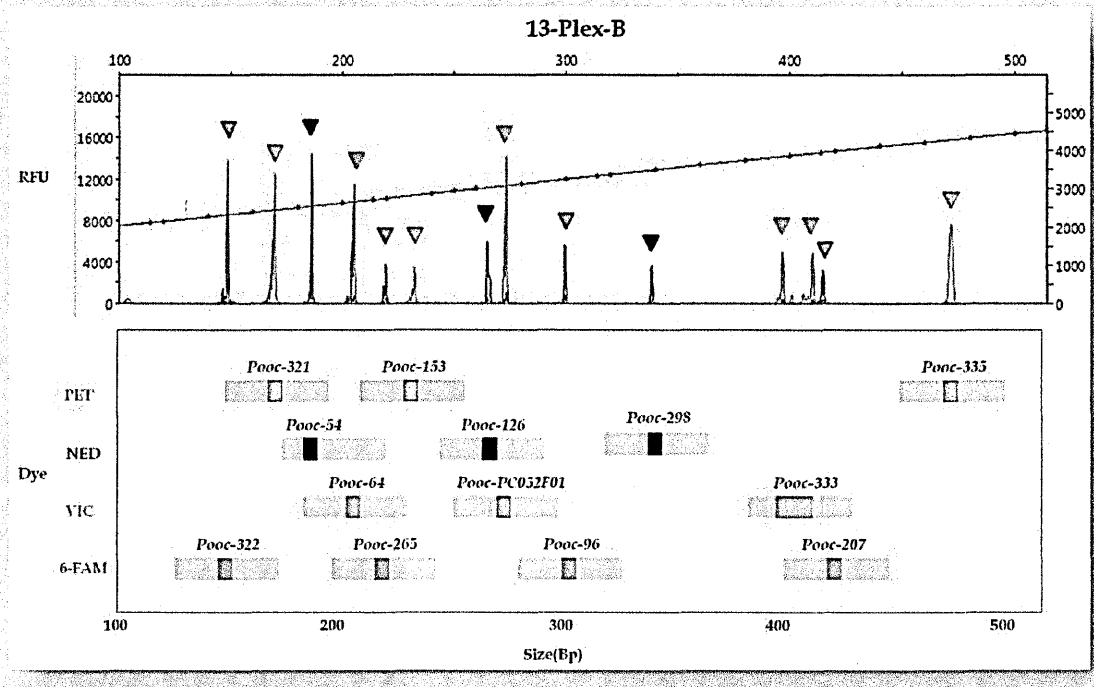
Welch, D.M. and M. Meselson. (2000). Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science*, 288: 1211–1215.

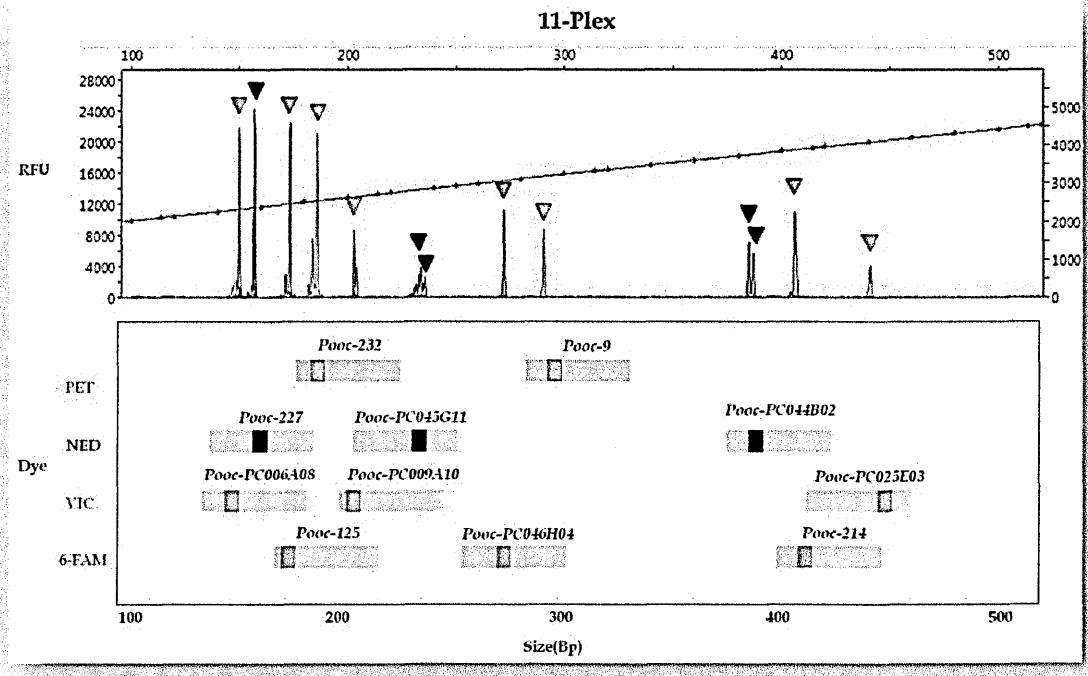
Wissler, L., Dattolo, E., Moore, a D., Reusch, T. B. H., Olsen, J. L., Migliaccio, M., Bornberg-Bauer, E., Procaccini, G. (2009). Dr. Zompo: an online data repository for *Zostera marina* and *Posidonia oceanica* ESTs. *Database : the journal of biological databases and curation*, bap009. doi:10.1093/database/bap009

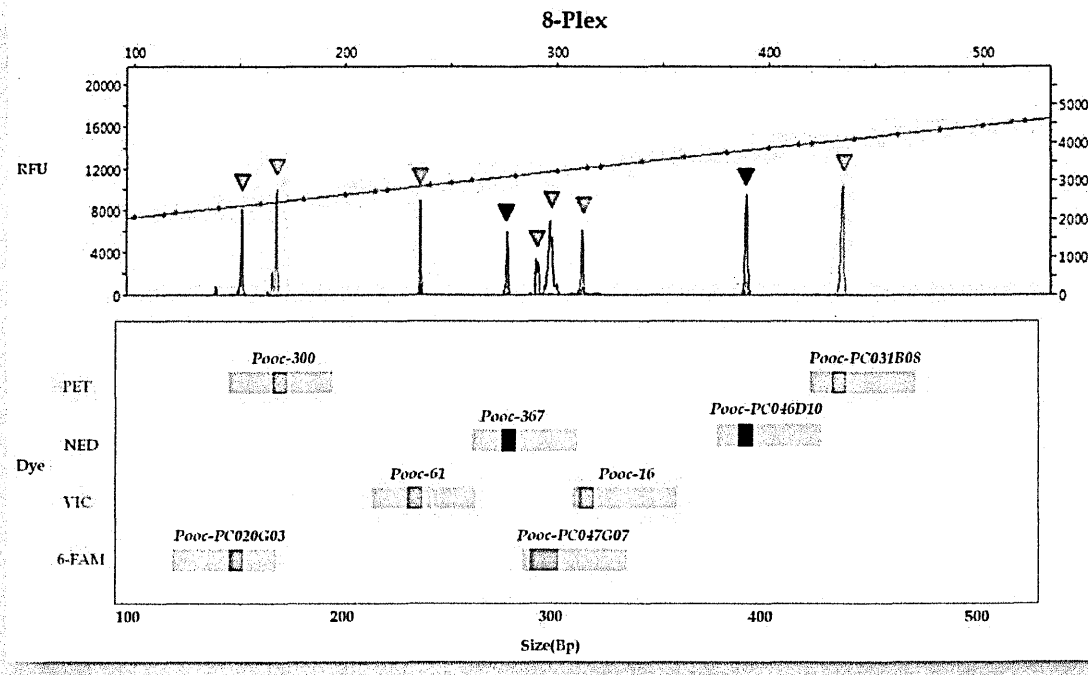
Zane, L., Bargelloni, L., Patarnello, T. (2002). Strategies for microsatellite isolation: a review. *Molecular ecology*, 11(1): 1–16.

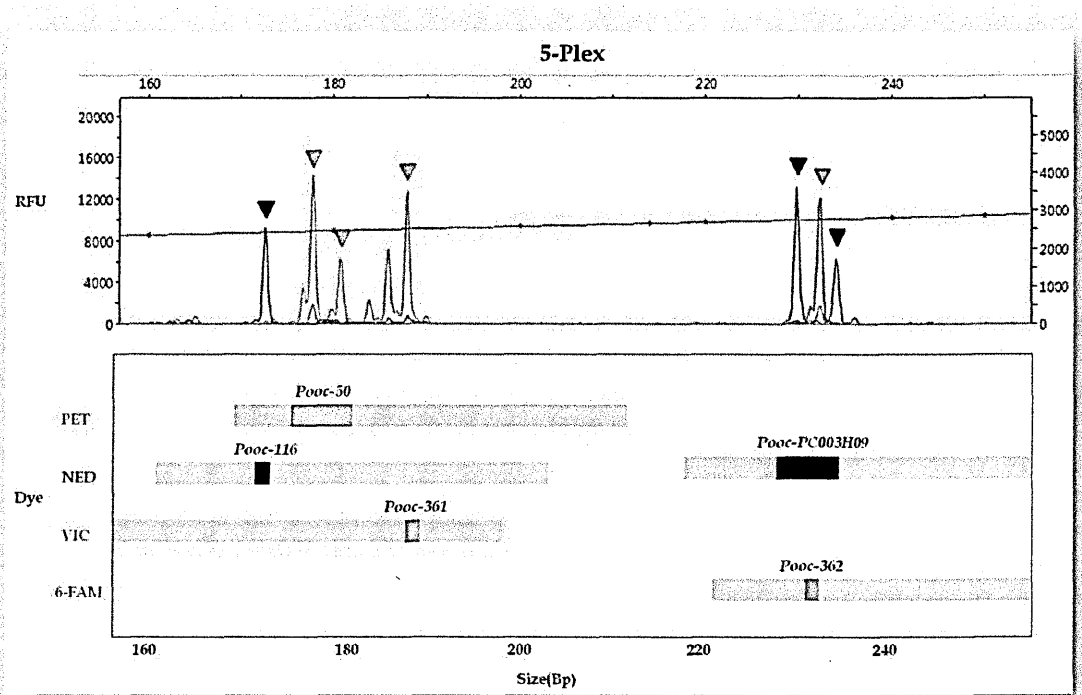
2.7-Appendix

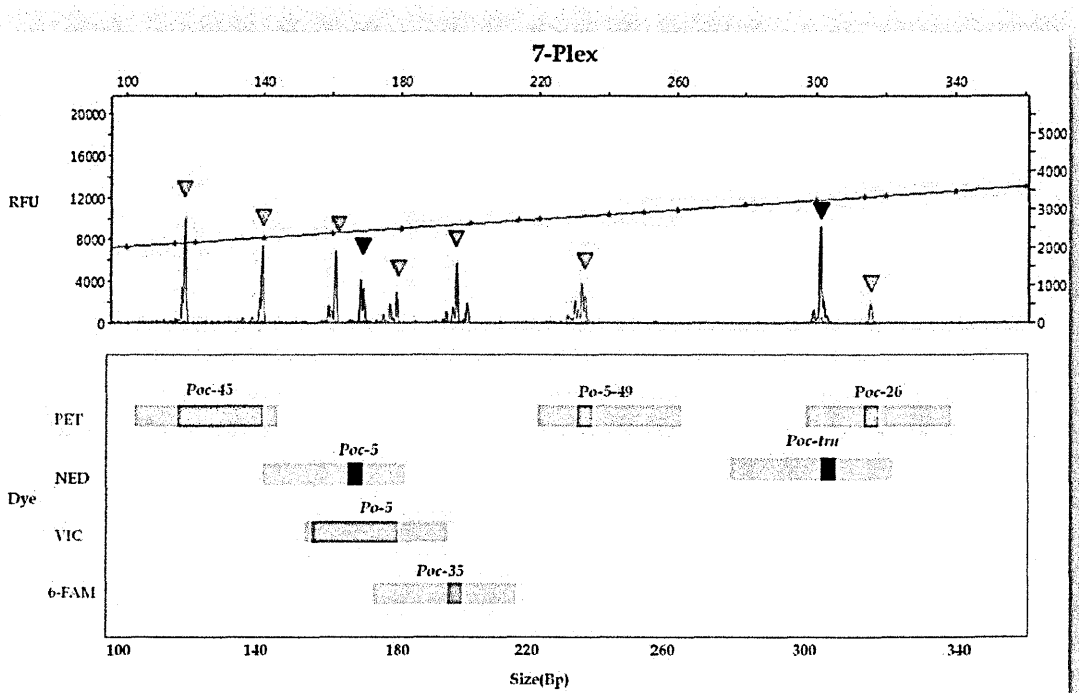
Figure 2.A1. Example of a typical electrophoretic profile obtained for the remaining 6 multiplexed PCR for one individual (upper panel) and diagram showing allele size range and fluorescent dyes of each of the multiplexed loci (lower panel). In upper panel, triangles indicate alleles at each locus. In lower panel, dark rectangles represent the observed allele range (base pair); light rectangles represent an arbitrary potential allele size range used during the multiplex design to avoid allele overlap of loci labeled with the same fluorescent dye.

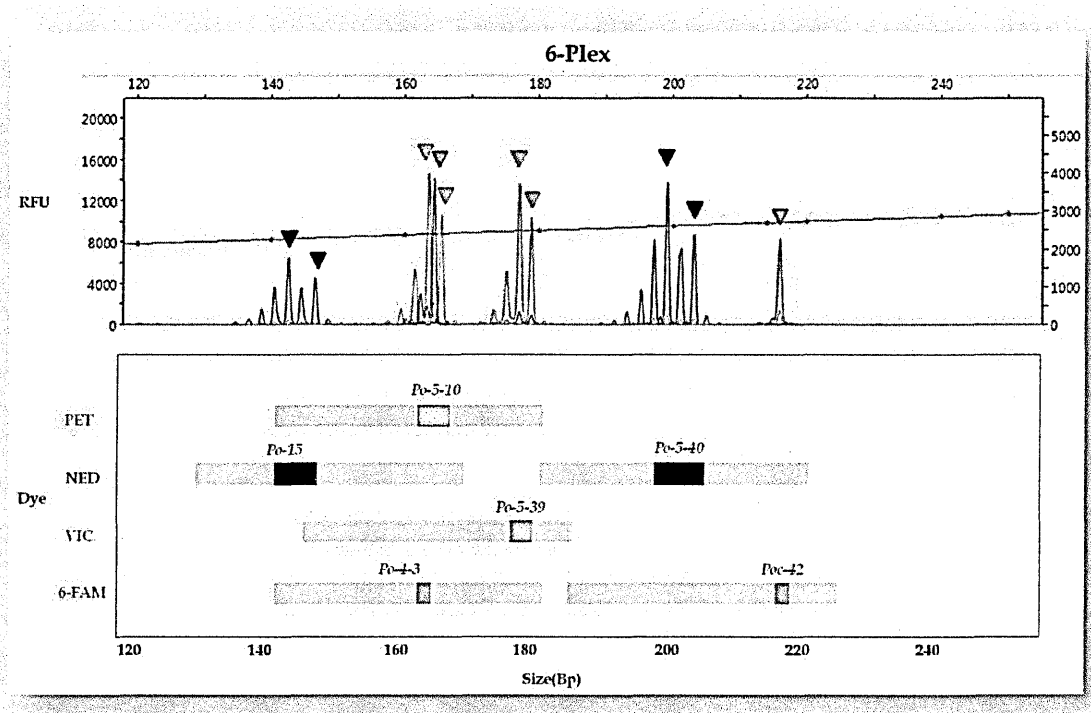



11-Plex
RFU
Dye
PET
NED
VIC
6-FAM
Pooc-232
Pooc-9
Pooc-227
Pooc-PC045G11
Pooc-PC044B02
Pooc-PC006A08
Pooc-PC009A10
Pooc-PC025E03
Pooc-125
Pooc-PC046H04
Pooc-214
Size(Bp)

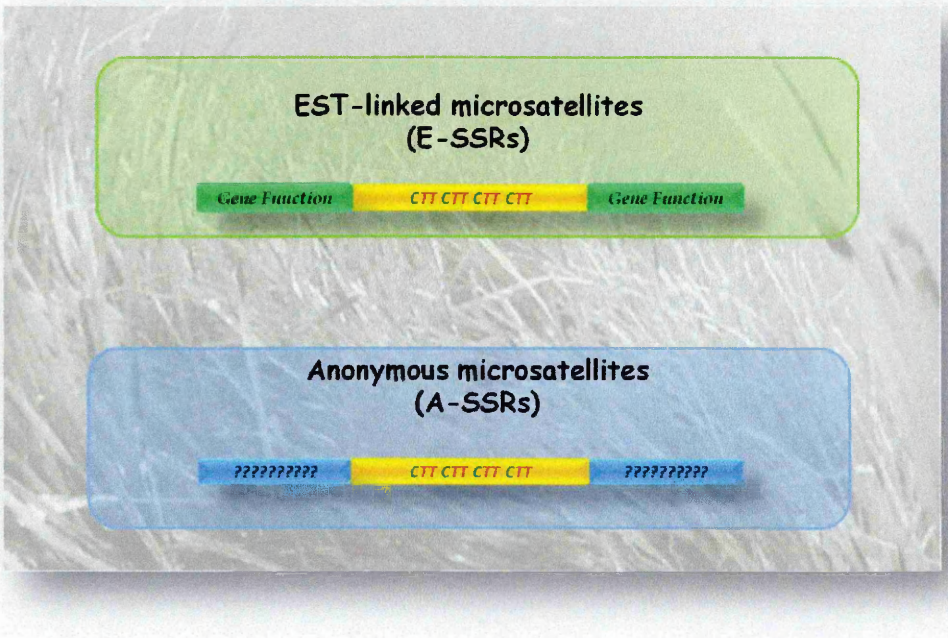

8-Plex
RFU
Dye
PET
NED
VIC
6-FAM
Pooc-300
Pooc-PC031B08
Pooc-367
Pooc-PC046D10
Pooc-61
Pooc-16
Pooc-PC020G03
Pooc-PC047G07
Size(Bp)


5-Plex
RFU
Dye
PET
NED
VIC
6-FAM
Pooc-50
Pooc-116
Pooc-361
Pooc-PC003H09
Pooc-362
Size(Bp)


7-Plex
RFU
Dye
PET
NED
VIC
6-FAM
Poc-45
Po-5-49
Poc-26
Poc-5
Poc-trn
Po-5
Poc-35
Size(Bp)


6-Plex
RFU
Dye
PET
NED
VIC
6-FAM
Po-5-10
Po-15
Po-5-40
Po-5-39
Po-4-3
Po-42
Size(Bp)

Chapter 3


EST-linked microsatellites
(E-SSRs)
Gene Function
CTT CTT CTT CTT
Gene Function
Anonymous microsatellites
(A-SSRs)
??????????
CTT CTT CTT CTT
??????????

3-Comparative performance of EST-linked and anonymous microsatellite loci in assessing genetic diversity in the seagrass *Posidonia oceanica*

3.1-Abstract

Anonymous and EST-linked microsatellites have been used to measure genetic variation in the wild. Due to their linkage with functional genes, EST-linked markers may show lower polymorphism and consequently lower resolution in the assessment of clonal diversity and population differentiation. Here, the performance of these two types of markers was compared in populations of *Posidonia oceanica* analyzed with 13 anonymous and 51 EST-linked microsatellites (E-SSRs). Anonymous microsatellites showed significantly higher levels of allelic diversity than EST-linked microsatellites, higher power in the analysis of clonality as well as in the study of population differentiation and cluster analysis. Conversely, E-SSRs showed overall lower variation but have the major advantage of leading to a better understanding of genes underlying local adaptation.

Keywords: anonymous microsatellites, expressed sequence tag-microsatellites, genetic diversity, polymorphism, marker comparison.

3.2-Introduction

Genetic diversity plays an important role in species persistence. Therefore, the analysis of the levels of genetic variation in natural populations is critical to the understanding of species evolution and to define appropriate strategies for the conservation and management of biodiversity. Several approaches have been adopted in the last 20 years to study organismal genetic diversity at population level (e.g. Agarwal et al., 2008, Finger and Klank, 2010). DNA approaches commonly employed in plant population genetics are based on AFLPs (amplified fragment length polymorphisms, Vos et al., 1995), RAPDs (randomly amplified polymorphic DNAs, Williams et al., 1990), ISSRs (inter simple sequence repeats, Zietkiewicz et al., 1994) and microsatellites (or SSRs, Tautz, 1989). Among those, the most commonly used markers in population genetics are microsatellites. Microsatellites are codominant hypervariable markers (e.g. Oliveira et al., 2006; Selkoe and Toonen, 2006; Kalia et al., 2011), especially suited to indentify genetic variation at population and individual levels. Among their limitations there are: 1) a species-specific marker development phase required (Squirrell et al. , 2003), 2) possible variation in the primer sites that can lead to null alleles (Pemberton et al., 1995) and 3) homoplasy due to identity between alleles by chance rather than descent (Goldstein and Schlötterer, 1999). Traditionally, only genomic microsatellites were identified from the genome of non model species which were not suited for functional studies. With the increasing availability of genomic tools for genetic model and non-model species, an increasing amount of sequence information is becoming available in public and private databases. These databases can be "mined" for SSR loci (Scott et al., 2000; Arnold et al., 2002; Eujayl

et al., 2004; Gao et al., 2004) that can be used in functional studies, because of their linkage with expressed sequence tags (Varshney et al., 2005).

As a consequence of this increasing availability, E-SSRs are used to study adaptive evolution in natural populations. The ability to indentify signatures of selection on these microsatellite loci depends on the strength of linkage from the locus under selection and may decay with recombination events (Wiehe, 1998). Therefore, the capability to indentify selective sweeps in loci linked to genes under directional selection depends on the "age" and strength of the selection event (Vigouroux et al., 2002). Given their linkage to genes under selection ESTs-linked markers are expected to be more conserved as compared to markers located in non-transcribed regions. Functional constraints in transcribed regions of the genome can result in a lower mutation rate both in the repeat region and in the priming site (Metzgar et al., 2000; Li et al., 2002). This should also reduce the presence of null alleles, that could occur when one or both primers extend across a splicing site or when mutations (*in-dels* or substitutions) are in the primer binding site (Lehman et al., 1996). In addition, as consequence of their lower mutation rate, E-SSRs have a higher level of transferability to related species than genomic SSRs, allowing a greater cross amplification of primer sets in range-wide studies (Varshney et al., 2005).

Repeat number and total length of SSRs in transcribed regions are smaller (Thiel et al., 2003; Buonaccorsi et al., 2012) as compared to genomic SSRs. In this shorter repeat array, the slippage of the DNA polymerase during PCR is reduced resulting in cleaner fragment profiles with fewer stutter bands for EST-SSRs than that typically observed from genomic SSRs. On the other hand, however, the lower

polymorphism in EST-SSRs as compared to genomic microsatellites, reduces sensitivity of the former in detecting within population variation (Varshney et al., 2005).

So far, direct comparisons of the performance of EST-SSRs (E-SSRs) and genomic SSRs (A-SSRs) in inferring population genetic structure have been reported in few studies. In plants, Woodhead et al. (2005) showed that, in the alpine lady-fern, E-SSRs are as efficient as neutral markers in the analysis of population genetic structure. Analyses of the difference of genomic SSR and E-SSR in assessing the genetic diversity of poplar clones (Song et al., 2012) revealed that genomic SSRs were more polymorphic than E-SSRs, but the latter were better at identifying distinct genotypes. Conversely, a study performed in Durum wheat (Eujayl et al., 2001) showed that the E-SSRs were less polymorphic but also less powerful in discriminating genotypes compared to the genomic SSRs. In corn rootworms, Kim et al. (2008) reported that genomic and E-SSRs performed similarly in estimating genetic diversity, population assignment and detection of distinct populations, while in lettuce (Simko et al. 2009) and in cucumber (Hu et al., 2011) anonymous SSRs showed higher polymorphisms than E-SSRs.

Studies performed in fish, generally agree on showing that E-SSRs are less polymorphic as compared to genomic SSRs (Coulibaly et al. 2005; Vilas et al. 2010; Buonaccorsi et al. 2012) with the exception of a flatfish species, *Senegalese sole*, where Molina-Luzón et al. (2012) found no significant differences between A-SSRs and E-SSRs in terms of number of alleles and observed and expected heterozygosities.

Overall, existing studies agree on the fact that A-SSRs are more polymorphic than E-SSRs, although there is not a common trend on which of the two types of markers is more informative in population genetic analyses.

In this study, the performance and resolution power of the new E-SSR loci presented in chapter 2 and A-SSR loci previously developed in *Posidonia oceanica* (Procaccini and Waycott, 1998; Alberto et al., 2003) were investigated in of six populations from the Mediterranean basin, with the aim of identifying the most informative marker type for future population genetic analyses. We expect to find a major polymorphism at the anonymous markers in comparison to E-SSRs markers that, being located in transcribed regions, are supposed to have a higher level of sequence conservation. Hence, we also expect that anomymous markers will be more powerful in the genotype discrimination as well in the analysis of population structure.

3.3-Materials and Methods

3.3.1-Collection of samples and DNA isolation

Six *Posidonia oceanica* populations (N= 40 to 53) were analyzed in this study (Figure 3.1). Each population was sampled at two different depths (-5 m and -20 m).

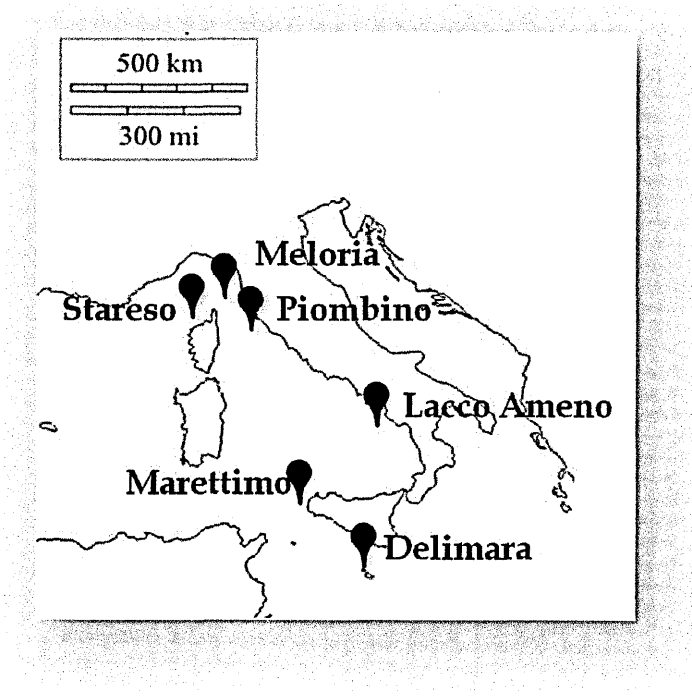


Figure 3.1. Geographical position of the *Posidonia oceanica* populations sampled. Delimara - Strait of Sicily, Malta (35°51'28"N; 14°33'07"E); Lacco Ameno - Central Tyrrhenian Sea, Italy (40°45'52''N; 13°53'29''E); Marettimo - Southern Tyrrhenian Sea (37°58'10"N; 12°04'76"E), Italy; Meloria - North Tyrrhenian Sea, Italy (43°31'41"N; 10°10'32"E); Piombino -North Tyrrhenian Sea, Italy (42°56'92"N; 10°38'99"E); Stareso - Corsica Sea, France (8°45'E; 42°35'N).

For Lacco Ameno, Marettimo, Piombino and Stareso populations, leaf tissues was cleaned of epiphytes and stored in 15 ml Falcon tubes (Sigma-Aldrich) filled with silica gels crystals (AppliChem). 30 mg of tissue were grounded in the TissueLyser MixerMill (Qiagen) and transferred in 96-well plates where DNA extraction was performed using the NucleoSpin® 96 Plant II kit (Macherey-Nagel), following the instructions of the manufacturer modified as in Migliaccio et al. (2005). For Delimara and Meloria populations, a hexadecyltrimethyl ammonium bromide (CTAB) extraction procedure was performed (see Material and Methods, Chapter 2).

3.3.2-Microsatellites selection

Two microsatellites types were analyzed in this study: 13 anonymous or genomic microsatellites (A-SSRs, Procaccini and Waycott, 1998; Alberto et al., 2003) and 51 EST-linked microsatellites (E-SSRs, see Chapter 2 for E-SSRs details).

3.3.3-Clonality and genetic diversity indices

To investigate the performance of the two classes of markers, clonal diversity and genetic diversity indices were calculated separately for each microsatellite class.
Before calculating population genetic parameters, identical genotypes were identified and excluded. Levels of clonality were estimated by calculating genotypic richness (R), defined as

$$R = \frac{G - 1}{N - 1}$$

where G is the number of distinct multilocus genotypes and N is the number of samples in the population.

Discriminating power (D) estimated for each microsatellite locus was used as a quantification tool to measure the efficiency of a given marker for the discrimination of genotypes. Following Tessier at al., 1999, if C is the confusion probability, i.e. the probability that two randomly chosen individuals from the sample have identical allelic patterns, then D=1-C represents the probability that two randomly chosen individuals have different patterns, and thus are distinguishable from one another. For the jth primer, the confusion probability C_j is equal to the sum of the different c_j for all I patterns generated by the primer:

$$C_j = \sum_{i=1}^{I} c_i = \sum_{i=1}^{I} p_i \frac{(Np_i - 1)}{N - 1}$$

Thus, the discriminating power of the jth primer is equal to:

$$D_j = 1 - C_j = 1 - \sum_{i=1}^{I} p_i \frac{(Np_i - 1)}{N - 1}$$

In order to select the best markers combination, including A-SSRs and E-SSRs, to use in future studies of clonal diversity in *P. oceanica*, the resolution power of the combined set of markers was also analyzed (see Appendix).

Number of alleles (N_a), allelic richness (A), expected (H_e) and observed (H_o) heterozygosity were calculated using the software FSTAT ver. 2.9.3 (Goudet, 2001). Private alleles were scored with Genalex ver. 6.5 (Peakall and Smouse, 2006).

The repeat number and the variance in the repeat number at each locus were calculated with the software Microsatellite analyzer (MSA) ver. 4.05 (Dieringer and Schlötterer, 2003).

The Polymorphism Information Content (PIC) was used as a tool to measure the information that a given marker locus could provide for the pool of genotypes. PIC values were calculated for estimates of marker informativeness according to the equation of Botstein et al. (1980) using the Excel Microsatellite Toolkit 3.1.1 (Park, 2001).

3.3.4-Population genetic structure analysis

F-statistics and population genetic structure analyses were performed separately for each microsatellite class.

F-statistics (F_{IS}, F_{IT} and F_{ST}) were calculated excluding identical clones within population using the software FSTAT (Goudet 2001).

To visualize the relationships among populations a PCoA (Principal Coordinate Analysis) was performed using GenAlEx ver. 6.5 (Peakall and Smouse, 2006).

Bayesian clustering method implemented in the software BAPS ver. 5.4 (Tang et al., 2009) was used to investigate the genetic structure of *Posidonia oceanica* populations. Individual clustering was estimated using K-values ranging from 1 to

21 (six populations sampled at two depths). 100 iterations of each K-value were used to judge the consistency of the simulation results. In each simulation, 200 reference individuals/population and 10 iterations were used to estimate the admixture coefficients of the reference individuals. Evidence for admixture was considered significant for individuals with P-values <0.05 (Corander and Marttinen, 2006).

3.4-Results

3.4.1-Comparison of A-SSRs and EST-SSRs in the assessment of clonality and genetic diversity indices

The 13 A-SSRs were polymorphic in all populations analyzed whereas of the 51 E-SSRs, 16 (31.37%) were polymorphic in at least one population, and 35 (68.63%) were always monomorphic (Figure 3.2).

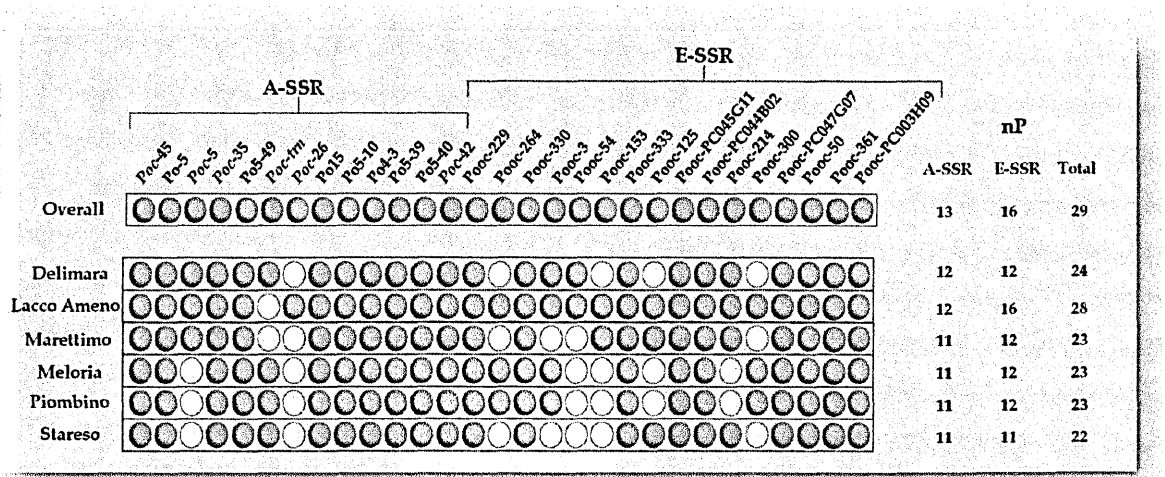


Figure 3.2. Visual representation of the screening for microsatellite polymorphism in the six populations analyzed and in the overall survey. Fill circles = polymorphic loci; empty circle = monomorphic loci; nP = number of polymorphic loci; A-SSRs = anonymous microsatellites; E-SSRs = EST-linked microsatellites, total = polymorphic loci including both A-SSRs and E-SSRs.

The number of alleles (N_a) in the A-SSRs ranged from 2 to 21 (mean A = 6.769 ± 5.890) and in the EST-SSR from 2 to 16 (mean A = 3.375 ± 3.462) (Table 3.1). Allelic richness (A) ranged from 2 to 14.917 (mean A = 5.218 ± 4.373) for A-SSRs and from 1.867 to 11.343 (mean A = 2.847 ± 2.313) for E-SSRs. Overall, the total

expected heterozygosity per locus (H_T) was 0.562 ± 0.275 (range 0.030-0.928), for A-SSRs and 0.382 ± 0.244 (range-0.101-0.880), for E-SSRs (Table 3.1).

Table 3.1. Characteristics of 13 polymorphic A-SSRs and 16 polymorphic E-SSRs in 6 populations of *Posidonia oceanica*. N_a = number of alleles; A = allelic richness; H_T = total expected heterozygosity; F_{IT}, F_{ST} and F_{IS} = F statistics; SD = standard deviation.

A-SSRs	Allele size	N_a	A	H_T	F_{IT}	F_{ST}	F_{IS}
Poc-45	116-140	2.000	2.000	0.501	-0.848	0.016	-0.878
Po-5	161-197	7.000	4.441	0.648	-0.031	0.206	-0.297
Poc-5	167-173	3.000	2.153	0.060	0.273	0.030	0.250
Poc-35	190-199	4.000	3.112	0.441	-0.225	0.155	-0.450
Po5-49	216-250	14.000	10.704	0.859	0.201	0.177	0.030
Poc-tm	291-315	4.000	3.035	0.551	-	-	-
Poc-26	309-316	2.000	1.552	0.030	0.002	0.081	-0.086
Po15	120-168	14.000	12.056	0.884	0.089	0.160	-0.085
Po5-10	151-165	5.000	3.648	0.588	0.279	0.179	0.122
Po4-3	160-170	6.000	4.48	0.694	0.199	0.302	-0.146
Po5-39	175-181	4.000	3.735	0.589	-0.008	0.155	-0.193
Po5-40	189-233	21.000	14.917	0.928	0.241	0.141	0.116
Poc-42	210-216	2.000	2.000	0.529	0.354	0.414	-0.102
Overall		6.769	5.218	0.562	0.044	0.168	-0.143
SD		5.890	4.373	0.275	0.326	0.109	0.300

continued

E-SSRs	Alleles size	N_a	A	Ht	F_{IT}	F_{ST}	F_{IS}
Pooc-229	150-168	3.000	3.000	0.594	0.161	0.092	0.076
Pooc-264	190-199	2.000	1.946	0.101	0.067	0.106	-0.044
Pooc-330	291-306	2.000	2.000	0.513	0.278	0.328	-0.074
Pooc-3	499-505	3.000	2.126	0.170	0.142	0.204	-0.077
Pooc-54	183-186	2.000	1.928	0.107	-0.019	0.176	-0.238
Pooc-153	232-246	2.000	1.966	0.122	-0.034	0.176	-0.256
Pooc-333	398-413	3.000	3.000	0.665	-0.032	0.160	-0.229
Pooc-125	173-187	2.000	1.986	0.147	-0.056	0.155	-0.250
Pooc-PC045G11	220-254	16.000	11.343	0.880	0.013	0.046	-0.035
Pooc-PC044B02	385-387	2.000	2.000	0.495	-0.796	0.001	-0.799
Pooc-214	406-408	2.000	1.867	0.078	-0.033	0.027	-0.062
Pooc-300	158-168	3.000	2.441	0.244	0.099	0.151	-0.062
Pooc-PC047G07	290-296	2.000	2.000	0.500	-0.977	0.001	-0.979
Pooc-50	178-181	2.000	2.000	0.500	-	-	-
Pooc-361	178-200	5.000	3.377	0.534	-0.160	0.163	-0.386
Pooc-PC003H09	230-236	3.000	2.575	0.459	-0.425	0.080	-0.550
Overall		3.375	2.847	0.382	-0.118	0.124	-0.264
SD		3.462	2.313	0.244	0.351	0.087	0.301

Population level descriptive statistics per population and separately for the two classes of markers is reported in Table 3.2. At the population level, the total number of alleles ranged from 34 (Stareso) to 56 (Delimara) and from 31 (Stareso) to 43 (Lacco Ameno) for the A-SSRs and the EST-SSRs, respectively. For A-SSRs, the number of alleles/locus ranged from 2 to 10 in Lacco, from 2 to 11 in Piombino, from 2 to 13 in Delimara and Marettimo, from 2 to 9 in Meloria and from 2 to 8 in Stareso. For E-SSRs, the number of alleles/locus ranged from 2 to 10 in Delimara and from 2 to 9 in Lacco Ameno, Marettimo, Meloria, Piombino and Stareso. Across populations, the largest mean number of alleles was shown by A-SSR locus Po5-40 (N_a across the 6 populations = 10.667) and by the EST-linked locus Pooc-PC045G11 (N_a across the 6 populations = 9.167).

Allelic richness ranged from 34 (Stareso) to 52.561 (Delimara), for A-SSR and from 30.512 (Stareso) to 42.291 (Lacco Ameno), for E-SSRS.

A-SSRs showed private alleles in all the 6 populations studied (Table 2). A total of 24 private alleles were observed, ranging from 1 (Piombino and Stareso) to 9 (Lacco Ameno). At locus level 11 (84.61%) out the 13 A-SSR loci, showed private alleles, with the locus Po5-49 showing the highest number of private alleles across populations (5 private alleles). E-SSRs showed a lower number of private alleles (five in total), ranging from 1 (Lacco Ameno) to 3 (Delimara) and were observed only in 3 southern populations (Delimara, Lacco Ameno and Meloria). Four (25%) out 16 polymorphic E-SSRs showed private alleles, with Pooc-361 showing the highest number (2 private alleles).

Table 3.2. Characteristics of E-SSRs and A-SSRs for the 6 populations analyzed. N_a = number of alleles; A = allelic richness, N_{ap} = number of private alleles; ; SD = standard deviation referred to N_a, allelic richness and N_{ap}; m = monomorphic locus; De = Delimara; La = Lacco Ameno; Ma = Marettimo; Me = Meloria; Pi = Piombino and St = Stareso

Locus	Na							Allelic richness						Nap					
A-SSR	De	La	Ma	Me	Pi	St	Across 6 populations	De	La	Ma	Me	Pi	St	De	La	Ma	Me	Pi	St
Poc-45	2	2	2	2	2	2	2.000	2.000	2.000	2.000	2.000	2.000	2.000	—	—	—	—	—	—
Po-5	4	4	4	3	3	2	3.333	3.385	3.969	3.579	3.000	3.000	2.000	1	1	—	—	—	—
Poc-5	2	2	2	m	m	m	1.500	1.999	1.999	1.983	m	m	m	—	—	1	—	—	—
Poc-35	2	3	4	3	3	3	3.000	2.000	2.999	3.628	2.978	2.459	3.000	—	—	1	—	—	—
Po5-49	8	9	8	6	8	3	7.000	7.841	8.963	7.585	5.950	7.433	3.000	1	2	2	—	—	—
Poc-trn	2	m	m	2	2	3	1.833	2.000	m	m	2.000	2.000	3.000	1	—	—	—	—	1
Poc-26	m	2	m	m	m	m	1.167	m	2.000	m	m	m	m	—	1	—	—	—	—
Po15	10	8	7	8	7	4	7.333	8.761	7.813	6.863	7.130	6.389	4.000	—	2	—	1	—	—
Po5-10	3	4	3	2	4	3	3.167	2.908	3.999	3.000	2.000	3.712	3.000	—	—	—	—	1	—
Po4-3	5	2	2	4	3	2	3.000	4.293	2.000	2.000	4.000	3.000	2.000	2	—	—	1	—	—
Po5-39	3	3	2	3	2	2	2.500	3.000	3.000	2.000	2.711	2.000	2.000	—	1	—	—	—	—
Po5-40	13	10	13	9	11	8	10.667	12.376	9.570	11.055	7.761	10.012	8.000	4	—	—	—	—	—
Poc-42	2	2	2	2	2	2	2.000	1.998	2.000	2.000	1.919	2.000	2.000	—	—	—	—	—	—
Total	56	51	49	44	47	34	48.500	52.561	50.312	45.693	41.449	44.005	34.000	9	7	4	2	1	1
Mean	4.67	4.25	4.45	4.00	4.27	3.09	3.731	4.380	4.193	4.154	3.768	4.000	3.091	1.80	1.40	1.33	1.00	1.00	1.00
SD	3.701	2.989	3.532	2.530	3.036	1.758	2.828	3.418	2.888	3.032	2.172	2.722	1.758	1.304	0.548	0.577	0.000	—	—

continued

Locus	Na							Allelic richness						Nap					
E-SSR	De	La	Ma	Me	Pi	St	Across 6 populations	De	La	Ma	Me	Pi	St	De	La	Ma	Me	Pi	St
Pooc-229	3	3	3	2	2	2	2.500	3.000	3.000	3.000	2.000	2.000	2.000	—	—	—	—	—	—
Pooc-264	m	2	m	2	2	m	1.500	m	2.000	m	1.774	2.000	m	—	—	—	—	—	—
Pooc-330	2	2	2	2	2	2	2.000	2.000	2.000	2.000	2.000	2.000	1.982	—	—	—	—	—	—
Pooc-3	2	2	m	2	2	m	1.667	1.960	1.940	m	2.000	2.000	m	1	—	—	—	—	—
Pooc-54	2	2	m	m	m	m	1.333	2.000	1.999	m	m	m	m	—	—	—	—	—	—
Pooc-153	m	2	2	m	m	m	1.333	m	2.000	2.000	m	m	m	—	—	—	—	—	—
Pooc-333	3	3	3	3	3	3	3.000	2.995	3.000	3.000	3.000	2.750	3.000	—	—	—	—	—	—
Pooc-125	m	2	2	m	m	2	1.500	m	2.000	2.000	m	m	1.999	—	—	—	—	—	—
Pooc-PC045G11	10	9	9	9	9	9	9.167	9.765	8.617	8.569	8.956	8.190	8.620	—	—	1	—	—	—
Pooc-PC044B02	2	2	2	2	2	2	2.000	2.000	2.000	2.000	2.000	2.000	2.000	—	—	—	—	—	—
Pooc-214	2	2	2	m	m	2	1.667	2.000	1.987	2.000	m	m	1.929	—	—	—	—	—	—
Pooc-300	m	3	m	2	2	m	1.667	m	2.998	m	2.000	2.000	m	—	1	—	—	—	—
Pooc-PC047G07	2	2	2	2	2	2	2.000	2.000	2.000	2.000	2.000	2.000	2.000	—	—	—	—	—	—
Pooc-50	2	2	2	2	2	2	2.000	2.000	2.000	2.000	2.000	2.000	2.000	—	—	—	—	—	—
Pooc-361	4	2	3	3	3	3	3.000	3.919	2.000	3.000	2.774	2.750	2.982	2	—	—	—	—	—
Pooc-PC003H09	2	3	3	3	2	2	2.500	2.000	2.750	3.000	2.774	2.000	2.000	—	—	—	—	—	—
Total	36	43	35	34	33	31	38.833	35.639	42.291	34.569	33.278	31.690	30.512	3	1	1	0	0	0
Mean	3	2.69	2.92	2.83	2.75	2.82	2.427	2.970	2.643	2.881	2.773	2.641	2.774	1.5	1	1	—	—	—
SD	2.296	1.740	1.975	1.992	2.006	2.089	1.874	2.230	1.648	1.855	1.988	1.771	1.980	0.707	—	—	—	—	—

Considering the allele sizes detected, it was observed that also the length in repeats number was different between the two classes of markers (Figure 3.3). Overall, the A-SSR sequence have very long repeat arrays (median = 14, range 3-40) in comparison with short repeat arrays in the E-SSRs (median = 7.5, range = 3-32) (Figure 3). For the A-SSRs, the minimum repeat was 3 for the locus *Poc*-26 in Lacco Ameno and the maximum repeat was 40 for the locus Po5-40 in Delimara. The median repeat number values per population ranged from 12.5 (Lacco Ameno) to 16 in Stareso for the A-SSRs. For E-SSRs, the minimum repeats was 3 for the locus Pooc-264 in Lacco Ameno, Meloria, Piombino and for the locus Pooc-54 in Delimara and Lacco Ameno. The maximum repeat was 32 for the locus Pooc-PC045G11 in Marettimo (Figure 3.3). Median repeat number values per population ranged from 6 in Delimara to 8 in Marettimo for E-SSRs.

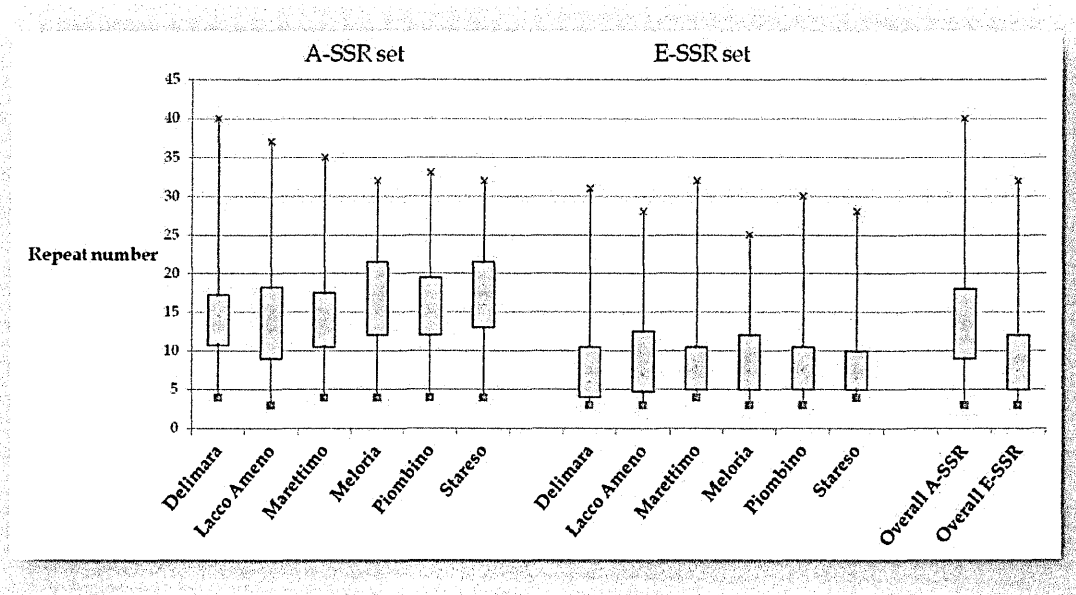


Figure 3.3. Boxplots of repeat number across loci analyzed overall and at population level. Median is showed within the boxes. Box limits indicate interquartile range among the loci.

The high variability observed for A-SSRs could be related to the repeat number or repeat length. The median number of repeats for anonymous microsatellites ranged from 3.5 (*Poc*-26) to 29 (Po5-40) for A-SSRs and from 3.5 (Pooc-54) to 23.5 (Pooc-PC045G11) for the E-SSRs (Figure 3.4).

Among the five different classes of microsatellites present in our sets (perfect di-nucleotides, perfect tri-nucleotides, composite tri-nucleotides, interrupted di-nucleotide and eptanucleotide), perfect di-nucleotide showed higher number of repeats both for A-SSR and E-SSR (Figure 3.4).

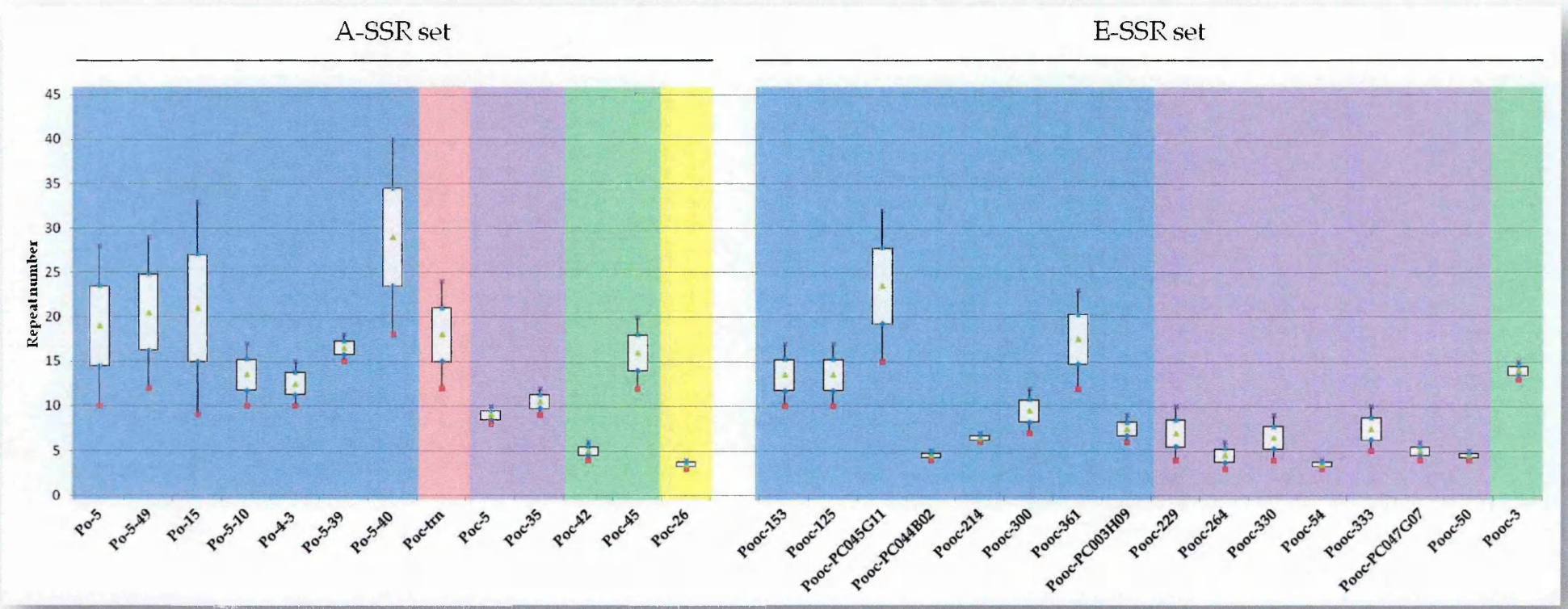


Figure 3.4. Boxplots of the number of repeats among the different classes of repeat motifs. In blue = perfect di-nucleotides; in purple = perfect tri-nucleotides; in yellow = epta-nucleotide; in green = composite tri-nucleotides; in pink = interrupted di-nucleotide.

The variance in the number repeats (Figure 3.5, overall) was higher for A-SSR (overall variance = 7.738 ± 10.841) and lower for E-SSRs (overall variance = 2.879 ± 3.845). The mean variance in the repeat number at population level for A-SSRs ranged from 5.552 (Meloria) to 13.290 (Lacco Ameno). E-SSRs showed a variance ranging from 1.999 (Meloria) to 4.31 (Marettimo). Stareso was the only population where variance was higher for E-SSRs than A-SSRs. For both the marker types an higher variance was observed for Delimara, Lacco Ameno and Marettimo populations in comparison to the northern population (Figure 3.5).

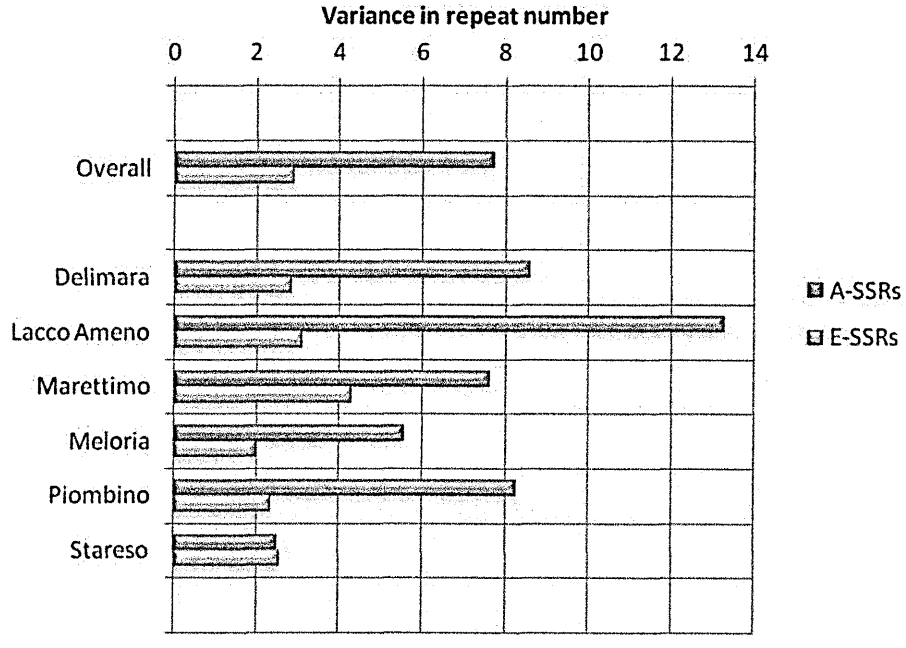


Figure 3.5. Variance in number repeats over loci analyzed overall and at population level.

In terms of observed and expected heterozygosity and Hardy-Weinberg equilibrium ($P < 0.05$) there were not significant differences between A-SSRs and E-SSRs (Table 3.3).

Table 3.3. Observed (H_o) and unbiased expected heterozygosities (uH_e) for the polymorphic A-SSRs and E-SSRs in the 6 populations of *Posidonia oceanica*. The cpDNA locus *Poc-trn* has not been used for the assessment of heterozygosities, Hardy-Weinberg equilibrium. De = Delimara; La = Lacco Ameno; Ma = Marettimo, Me = Meloria, Pi = Piombino, St = Stareso; m = monomorphic locus; SD = standard deviation; P = P-values of exact test of Hardy-Weinberg equilibrium (asterisks indicate significant departure from Hardy-Weinberg equilibrium, *=P<0.05).

	Ho						uHe						P					
A-SSRs																		
Locus	**De**	**La**	**Ma**	**Me**	**Pi**	**St**	**De**	**La**	**Ma**	**Me**	**Pi**	**St**	**De**	**La**	**Ma**	**Me**	**Pi**	**St**
***Poc*-45**	0.974	0.848	1.000	1.000	0.946	0.704	0.506	0.496	0.506	0.507	0.505	0.465	0.000*	0.000*	0.000*	0.000*	0.000*	0.009*
Po-5	0.513	0.879	0.535	0.605	0.595	0.963	0.436	0.659	0.452	0.498	0.537	0.509	0.197	0.014*	0.191	0.079	0.291	0.000*
***Poc*-5**	0.154	0.121	0.000	m	m	m	0.144	0.116	0.090	m	m	m	1.000	1.000	0.000*	m	m	m
***Poc*-35**	0.974	0.303	0.884	0.553	0.054	0.444	0.506	0.272	0.626	0.424	0.054	0.377	0.000*	1.000	0.000*	0.109	1.000	0.765
Po5-49	0.688	0.939	0.605	0.500	0.649	0.815	0.832	0.820	0.676	0.712	0.668	0.514	0.002*	0.063	0.042*	0.000*	0.049*	0.001*
Poc-trn	-	m	m	-	-	-	-	m	m	-	-	-	-	m	m	-	-	-
***Poc*-26**	m	0.182	m	m	m	m	m	0.168	m	m	m	m	m	1.000	m	m	m	m
Po15	0.872	0.788	0.837	0.895	0.919	0.444	0.806	0.814	0.795	0.762	0.729	0.517	0.010*	0.000*	0.013*	0.008*	0.004*	0.009*
Po5-10	0.436	0.455	0.326	0.289	0.541	0.556	0.475	0.450	0.617	0.321	0.540	0.461	0.142	0.891	0.000*	0.610	0.553	0.037*
Po4-3	0.333	0.333	0.767	0.447	0.629	0.815	0.313	0.357	0.488	0.672	0.540	0.509	0.342	0.649	0.000*	0.001*	0.105	0.002*
Po5-39	0.872	0.606	0.512	0.579	0.432	0.556	0.643	0.634	0.427	0.505	0.373	0.409	0.009*	0.009*	0.279	0.597	0.652	0.137
Po5-40	0.889	1.000	0.581	0.368	0.811	0.630	0.900	0.822	0.839	0.724	0.786	0.682	0.000*	0.022*	0.000*	0.000*	0.119	0.183
***Poc*-42**	0.128	0.485	0.357	0.053	0.324	0.889	0.122	0.485	0.492	0.052	0.275	0.509	1.000	1.000	0.111	1.000	0.561	0.000*
Mean	0.621	0.578	0.582	0.529	0.590	0.681	0.517	0.508	0.546	0.518	0.501	0.495						
SD	0.323	0.308	0.288	0.275	0.274	0.184	0.263	0.249	0.204	0.218	0.219	0.081						

continued

	Ho						uHe						P					
E-SSRs																		
Locus	**De**	**La**	**Ma**	**Me**	**Pi**	**St**	**De**	**La**	**Ma**	**Me**	**Pi**	**St**	**De**	**La**	**Ma**	**Me**	**Pi**	**St**
Pooc-229	0.615	0.656	0.643	0.452	0.531	0.121	0.627	0.530	0.632	0.444	0.496	0.507	0.854	0.121	0.722	1.000	0.729	0.000*
Pooc-264	m	0.219	m	0.032	0.313	m	m	0.198	m	0.032	0.310	m	m	1.000	m	1.000	1.000	m
Pooc-330	0.615	0.156	0.393	0.323	0.719	0.091	0.483	0.329	0.456	0.275	0.503	0.088	0.221	0.009*	0.671	0.568	0.030*	1.000
Pooc-3	0.040	0.063	m	0.483	0.281	m	0.040	0.062	m	0.460	0.246	m	1.000	1.000	m	1.000	1.000	m
Pooc-54	0.500	0.156	m	m	m	m	0.382	0.146	m	m	m	m	0.283	1.000	m	m	m	m
Pooc-153	m	0.500	0.259	m	m	m	m	0.381	0.230	m	m	m	m	0.147	1.000	m	m	m
Pooc-333	0.423	0.906	0.778	0.677	0.438	0.879	0.399	0.627	0.660	0.565	0.517	0.598	1.000	0.000*	0.054	0.272	0.358	0.000*
Pooc-125	m	0.563	0.214	m	m	0.152	m	0.411	0.195	m	m	0.142	m	0.070	1.000	m	m	1.000
Pooc-PC045G11	0.885	0.969	0.929	0.767	0.719	0.939	0.834	0.839	0.836	0.875	0.811	0.837	0.330	0.001*	0.000*	0.005*	0.044	0.418
Pooc-PC044B02	0.885	0.844	0.857	0.903	1.000	0.848	0.503	0.496	0.499	0.508	0.508	0.496	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
Pooc-214	0.115	0.094	0.214	m	m	0.061	0.111	0.091	0.195	m	m	0.060	1.000	1.000	1.000	m	m	1.000
Pooc-300	m	0.484	m	0.484	0.344	m	m	0.432	m	0.405	0.396	m	m	0.291	m	0.384	0.647	m
Pooc-PC047G07	0.917	1.000	1.000	1.000	1.000	1.000	0.507	0.508	0.509	0.508	0.508	0.508	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
Pooc-50	1.000	1.000	1.000	1.000	1.000	1.000	0.510	0.508	0.509	0.508	0.508	0.508	0.000*	0.000*	0.000*	0.000*	0.000*	0.000*
Pooc-361	0.308	0.563	0.536	0.323	1.000	0.939	0.278	0.437	0.551	0.362	0.523	0.551	1.000	0.121	1.000	0.677	0.000*	0.000*
Pooc-PC003H09	0.923	0.813	0.607	0.935	0.281	0.394	0.507	0.502	0.464	0.521	0.246	0.321	0.000*	0.000*	0.199	0.000*	0.000*	0.307
Mean	0.602	0.561	0.619	0.615	0.635	0.584	0.432	0.406	0.478	0.455	0.464	0.420						
SD	0.331	0.341	0.299	0.314	0.307	0.414	0.215	0.204	0.194	0.196	0.153	0.241						

In agreement with the heterozygosity values, A-SSR showed that all populations have negative F_{IS} values ranging from -0.387 (Stareso) to -0.022 (Meloria) (Table 3.4). E-SSRs showed negative F_{IS} values in all the populations analyzed ranging from -0.406 (Delimara) to -0.302 (Marettimo) and accounted for a major excess of heterozygosity compared with A-SSRs.

Table 3.4. F_{IS} for all the populations analyzed. m = monomorphic locus. * = significant at 99% confidence interval.

Locus	F_{IS} Per population :					
A-SSR	**De**	**La**	**Ma**	**Me**	**Pi**	**St**
Poc-45	-0.949	-0.730	-1.000	-1.000	-0.895	-0.529
Po-5	-0.180	-0.341	-0.187	-0.218	-0.108	-0.926
Poc-5	-0.070	-0.049	1.000	m	m	m
Poc-35	-0.949	-0.115	-0.418	-0.308	-0.007	-0.182
Po-5-49	0.176	-0.149	0.106	0.300	0.029	-0.604
Poc-trn	-	m	m	-	-	-
Poc-26	m	-0.085	m	m	m	m
Po-15	-0.083	0.033	-0.054	-0.176	-0.265	0.143
Po-5-10	0.084	-0.011	0.475	0.100	-0.001	-0.211
Po-4-3	-0.065	0.066	-0.582	0.337	-0.167	-0.620
Po-5-39	-0.362	0.045	-0.202	-0.149	-0.161	-0.368
Po-5-40	0.012	-0.220	0.309	0.494	-0.032	0.078
Poc-42	-0.056	0.000	0.276	-0.014	-0.180	-0.773
Total	-0.205*	-0.141*	-0.067*	-0.022*	-0.181*	-0.387*
E-SSR	**De**	**La**	**Ma**	**Me**	**Pi**	**St**
Pooc-229	0.018	-0.242	-0.018	-0.017	-0.073	0.764
Pooc-264	m	-0.107	m	0.000	-0.010	m
Pooc-330	-0.282	0.529	0.142	-0.176	-0.438	-0.032
Pooc-3	0.000	-0.016	M	-0.051	-0.148	m
Pooc-54	-0.316	-0.069	m	m	m	m
Pooc-153	m	-0.319	-0.130	m	m	m
Pooc-333	-0.062	-0.456	-0.182	-0.202	0.156	-0.480
Pooc-125	m	-0.378	-0.102	m	m	-0.067
Pooc-PC045G11	-0.062	-0.157	-0.113	0.126	0.115	-0.124
Pooc-PC044B02	-0.786	-0.722	-0.742	-0.803	-1.000	-0.730
Pooc-214	-0.042	-0.033	-0.102	m	m	-0.016
Pooc-300	m	-0.122	m	-0.200	0.135	m
Pooc-PC047G07	-0.840	-1.000	-1.000	-1.000	-1.000	-1.000
Pooc-50	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000
Pooc-361	-0.108	-0.295	0.029	0.111	-0.939	-0.725
Pooc-PC003H09	-0.852	-0.633	-0.315	-0.820	-0.148	-0.231
Total	-0.406*	-0.392*	-0.302*	-0.358*	-0.377*	-0.400*

Genotypic richness was assessed with the anonymous microsatellites (A-SSRs) and EST-linked microsatellites (E-SSRs), separately. Two levels of analysis were considered: the overall level (including the entire populations data set) and the population level.

Overall, A-SSRs revealed a higher genotypic richness (R) in comparison to E-SSRs (Table 3.5). Starting from the initial data set, consisting of 271 multilocus genotypes, the A-SSRs successfully discriminated 217 genotypes (R = 0.800) while the E-SSRs identified only 182 distinct genotypes (R = 0.670).

Table 3.5. Clonal diversity in the *Posidonia oceanica* populations studied. G = number of distinct multilocus genotypes; N= number of individuals; R (G-1/N1) = genotypic richness. Overall = analysis including the entire data set; SD = standard deviation.

	A-SSR set			E-SSR set		
	G	N	R	G	N	R
Overall	217	271	0.800	182	271	0.670
Delimara	39	40	0.974	26	340	0.641
Lacco Ameno	33	53	0.615	32	53	0.596
Marettimo	43	44	0.977	28	44	0.628
Meloria	38	40	0.949	31	40	0.769
Piombino	37	43	0.857	32	43	0.738
Stareso	27	51	0.520	33	51	0.640
Mean			0.815			0.669
SD			0.182			0.063

The analysis of clonal diversity performed at population level with the A-SSRs (Table 3.5) showed a genotypic richness (R) within populations ranging from 0.520

(Stareso) to 0.977 (Marettimo; mean R = 0.815 ± 0.182); E-SSRs showed a lower clonal diversity with R values ranging from 0.596 (Lacco Ameno) to 0.769 (Meloria; mean R = 0.669 ± 0.063; Table 3.5). In particular, A-SSRs markers alone seem to have a higher resolution in the discrimination of genetically distinct individuals in comparison to E-SSRs, except in Stareso, where the E-SSRs revealed higher genotypic richness in comparison to A-SSRs (genotypic richness ESSRs = 0.640 vs. genotypic richness ASSRs = 0.520; Figure 3.6).

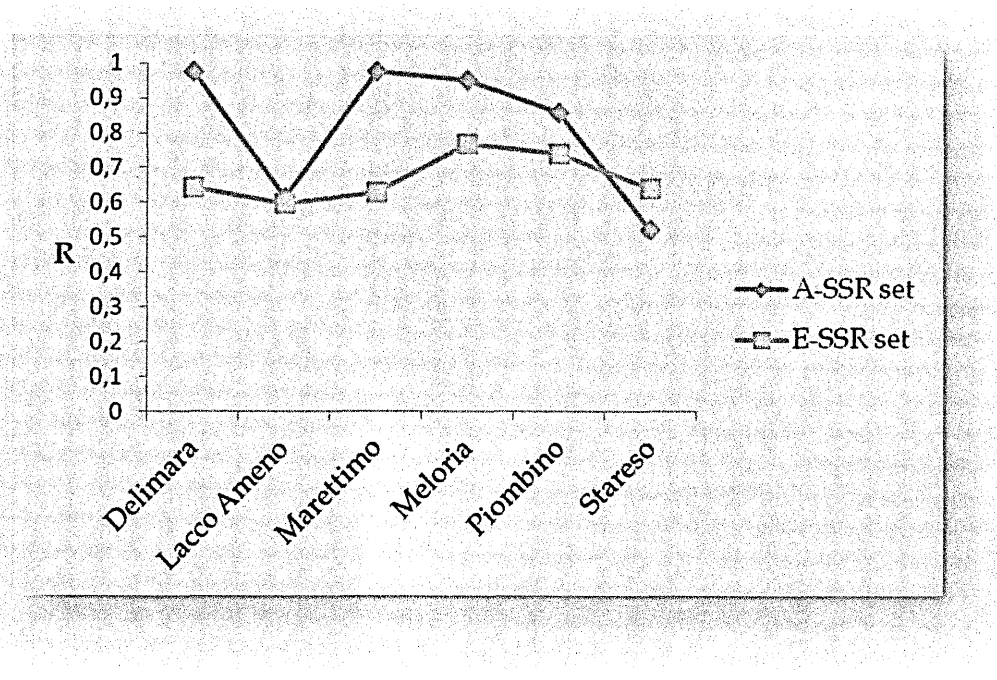


Figure 3.6. Genotypic richness at population level for the A-SSRs and E-SSRs. R = genotypic richness.

The discriminating power calculated for each polymorphic microsatellite locus is reported in Table 3.6. Overall, 13 A-SSRs and 16 E-SSRs were polymorphic and were used in this calculation (see Figure 3.2 for details). The number of patterns (or genotypes) identified from A-SSR ranged from 2 (*Poc*-26) to 70 (Po5-40; mean

number of patterns = 16.462 ± 21.620) and was higher than E-SSR, ranging from 1 (Pooc-50 is a polymorphic locus with two alleles always present in heterozygosis) to 51 (Pooc-PC045G11; mean number of patterns = 6.375 ± 12.027). Discriminating power (D) was higher for A-SSRs (mean D = 0.613 ± 0.318) than for E-SSRs (mean D = 0.394 ± 0.296). Conversely, confusion probability values were lower for A-SSR (mean C = 0.387 ± 0.318) and higher for E-SSRs (mean C = 0.606 ± 0.296). With respect to the discriminating power, microsatellites were ranked in classes from 1 (most discriminating locus) to 29 (less discriminating locus). Loci do not rank in D-classes coherently in accordance to their belonging to the two sets, with e.g., the EST-linked locus Pooc-PC045G11 having the highest discriminating power and the E-SSR Pooc-50 having the lowest (Table 3.6).

Table 3.6. Overall polymorphism of A-SSRs and E-SSRs. Number of patterns = number of genotypes; C = confusion probability; D = discriminating power; PIC = Polymorphic Information Content;. D class order in the score of D; PIC class order in the score of PIC

	Number of patterns	C	D	PIC	D-Class	PIC-class
A-SSRs						
***Poc*-45**	3	0.847	0.153	0.375	24	16
Po-5	11	0.223	0.777	0.570	7	6
***Poc*-5**	3	0.915	0.085	0.050	26	28
***Poc*-35**	5	0.373	0.627	0.384	14	13
Po-5-49	40	0.073	0.927	0.817	4	4
Poc-trn	4	0.501	0.499	0.397	16	12
***Poc*-26**	2	0.922	0.078	0.039	27	29
Po15	47	0.048	0.952	0.830	3	3
Po5-10	9	0.255	0.745	0.468	8	10
Po4-3	10	0.217	0.783	0.585	6	5
Po5-39	7	0.285	0.715	0.478	10	9
Po5-40	70	0.037	0.963	0.888	2	1
***Poc*-42**	3	0.340	0.660	0.372	12	18
Mean	16.462	0.387	0.613	0.481		
SD	21.620	0.318	0.318	0.265		

	Number of patterns	C	D	PIC	D-Class	PIC-class

E-SSRs						
Pooc-229	6	0.260	0.740	0.486	9	8
Pooc-264	3	0.850	0.150	0.078	25	17
Pooc-330	3	0.343	0.657	0.369	13	19
Pooc-3	4	0.612	0.388	0.130	17	22
Pooc-54	2	0.813	0.187	0.093	22	26
Pooc-153	2	0.796	0.204	0.102	20	24
Pooc-333	6	0.215	0.785	0.566	5	7
Pooc-125	2	0.757	0.243	0.122	19	23
Pooc-PC045G11	51	0.036	0.964	0.870	1	2
Pooc-PC044B02	3	0.819	0.181	0.373	23	17
Pooc-214	2	0.808	0.192	0.096	21	25
Pooc-300	5	0.615	0.385	0.218	18	21
Pooc-PC047G07	2	0.985	0.015	0.375	28	15
Pooc-50	1	1.000	0.000	0.375	29	14
Pooc-361	7	0.306	0.694	0.467	11	11
Pooc-PC003H09	3	0.486	0.514	0.363	15	20
Mean	6.375	0.606	0.394	0.318		
SD	12.027	0.296	0.296	0.219		

To have a qualitative assessment of the markers resolution, the PIC (Polymorphic Information Content) value was evaluated (Table 3.6). The PIC value of each SSR locus was determined by both the number of alleles and their frequency distribution in the complete population set and was used to assess their informativeness level. For ASSRs, the average PIC was 0.481 ± 0.265, ranging from 0.039 (*Poc*-26) to 0.888 (Po5-40). The E-SSRs showed a mean PIC value of 0.318 ± 0.219, with values ranging from 0.078 (Pooc-264) to 0.870 (Pooc-PC045G11). In conclusion, the informativeness of the EST-SSR markers appeared to be lower than A-SSR markers, although, following the criteria of Botstein et al. (1980), both sets of loci appeared to be only moderately informative (0.25 < PIC <0.5). With respect to the PIC values, microsatellites were ranked in classes from 1 (most informative locus) to 29 (less informative locus, Table 3.6). As for the order in D-classes, loci are not ranked in PIC-classes coherently in accordance to their

belonging to the two sets, with e.g., the A-SSR Po5-40 the most informative locus and the A-SSR *Poc*-26 the less informative. The comparison between the two order classes (D-class and PIC-class) showed a good correspondence for the first eleven loci.

Although the A-SSRs indicated higher discrimination in comparison to E-SSRs, the assessment of clonal diversity was also performed with the combined set of markers to investigate if the use of the two types of markers together allowed a better resolution that one set alone. This analysis, being beyond the scope of this study, is reported in the appendix nevertheless it provided important indication that the combination of the two classes of marker allowed to identify an efficient set of markers to be used in further studies of clonal diversity in *P. oceanica* (see Appendix for more details).

3.4.2-Comparison of anonymous versus EST-SSRs in the assessment of gene flow and partitioning of genetic variation

F_{ST} values for each locus analyzed at overall level are shown in Table 3.1. Average F_{ST} values across loci are 0.168 ± 0.109 (ranging from 0.016 to 0.414), for A-SSRs F_{ST}, and 0.124 ± 0.087 (ranging from 0.001 to 0.328), for E-SSRs.

Pairwise estimates among the 6 *Posidonia oceanica* populations were calculated based on the two types of SSR markers (Table 3.7). Pairwise F_{ST} estimates from A-SSRs were higher (F_{ST} range between population = 0.047-0.202) than the ones from E-SSRs (F_{ST} range between populations = 0.032-0.105).

Table 3.7. Population pairwise F_{ST} estimates from 13 genomic (first value) and16 EST- (second value) between 6 *Posidonia oceanica* populations

	Delimara	Lacco Ameno	Marettimo	Meloria	Piombino	Stareso
Delimara	0.000 0.000					
Lacco Ameno	0.069 0.059	0.000 0.000				
Marettimo	0.092 0.045	0.060 0.032	0.000 0.000			
Meloria	0.181 0.051	0.149 0.048	0.106 0.059	0.000 0.000		
Piombino	0.175 0.065	0.122 0.054	0.118 0.057	0.047 0.041	0.000 0.000	
Stareso	0.202 0.084	0.171 0.084	0.144 0.068	0.099 0.105	0.110 0.075	0.000 0.000

Principal Coordinate Analysis (PCoA), performed for genomic and EST-SSRs is presented in Figure 3.7. Also PCoA analysis revealed that A-SSRs have an higher power in resolution for genetic clustering. Principal coordinate 1 (x axis) and principal coordinate 2 (y axis) explain together the 61.71% of the total variation for the genomic SSRs and 48.35% for the EST-SSRs. The coordinate 1 for genomic SSRs clearly distinguished Meloria, Piombino and Stareso from the Delimara, Lacco Ameno and Marettimo samples, according to their geographical location. The coordinate 2 also separates Delimara from the other southern populations. The plot obtained from EST-SSRs does not give a clear separation among populations, not showing a clear genetic separation between southern and northern

populations The genomic set clearly better discriminated the 6 population groups than the E-SSRs.

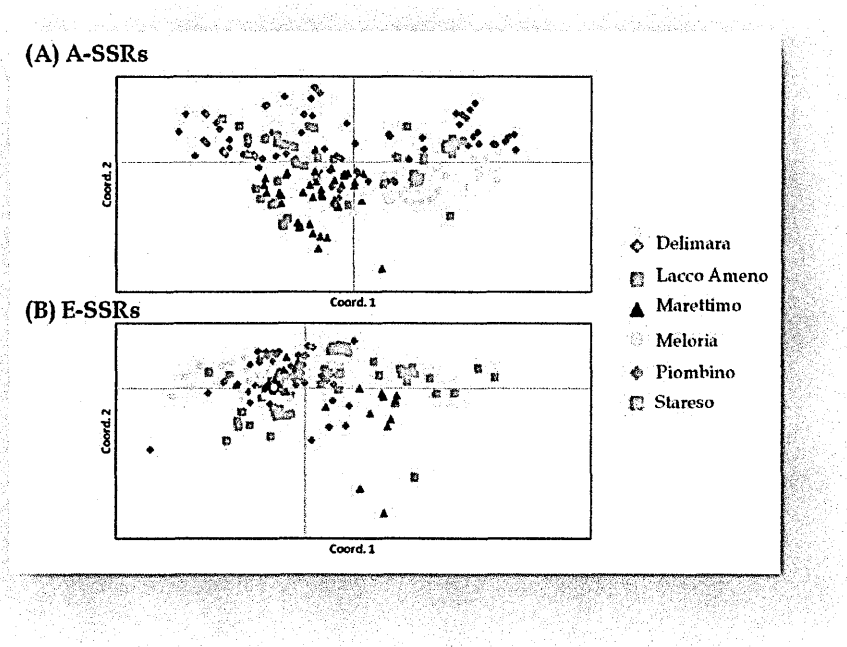


Figure 3.7. Principal coordinate plots of the distance matrices between individuals from *Posidonia oceanica* populations, calculated using (A) 13 genomic SSRs, axes explain 61.77% (46.53 + 15.18) of the total variation; (B) 16 EST-SSRs, axes explain 48.35% (29.39 + 18.96) of the total variation.

Cluster analysis with the best partitioning (k) is showed in Figure 3.8 for the two microsatellite sets. A-SSRs identify 15 genetic clusters for the individual level analysis whereas E-SSRs identified 9 independent clusters. Admixture is low in all samples and statistically significant in 3 of the 217 individual multilocus genotypes identified with A-SSRs (indicated with asterisks in Figure 3.8).

No significant admixture is revealed by E-SSRs analysis. The most striking pattern emerging from the Bayesian approach is the strong pattern of sub-structure within populations that is outlined in particular in the analysis with genomic microsatellites. This sub-structure of the populations could depend from the

sampling strategy according to which individuals were collected along a bathymetric gradient at two different depths and could reflect a differentiation between samples coming from the shallow and the deep sites of the meadow.

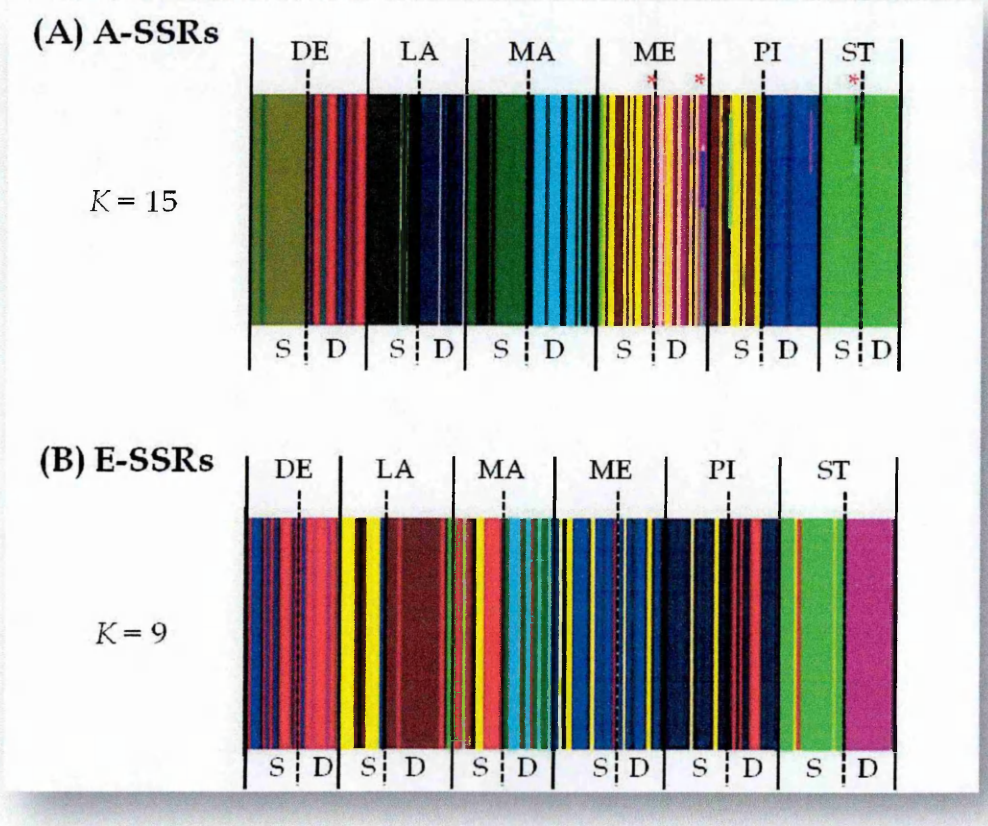


Figure 3.8. Inferred clusters of individual genotypes using A-SSRs (A) and E-SSRs (B). Shown are K-values at which clusters that had significant biological meaning were achieved. For each analysis is reported the K- value for which the highest likelihood was obtained. Each vertical bar represents one individual multilocus genotype. Individuals with multiple colors have admixed genotypes from multiple clusters. Statistically significant admixture for an individual multilocus genotype is noted with a red asterisk above. Each color represents the most likely ancestry of the cluster from which the genotype or partial genotype was derived. Clusters of individuals are represented by colors, populations are separated by black vertical lines and subpopulations (S = shallow subpopulation; D = deep subpopulation) are separated by dashed black vertical lines.

3.5-Discussion

A-SSR microsatellites showed a higher polymorphism than the E-SSR microsatellites. A-SSR, showed higher number of alleles in accordance with previous studies performed on other species, whereas E-SSRs were less polymorphic, likely because of their linkage with coding gene regions (Eujayl et al., 2001; Simko et al., 2009; Hu et al., 2011, Vilas et al.,2010, Buonaccorsi et al., 2012; Song et al., 2012).

In this study, the low diversity in EST-SSRs was associated with low repeat number and low variance in repeat number as compared to the same measures in anonymous SSRs. It was not surprising, because it is known that functional constraints may limit microsatellite repeat expansion in expressed portions of the genome (Metzgar et al. 2000) and the repeat number is probably the best characterized factor that affect the mutation rate of microsatellites (Chakraborty et al. 1997; Ellegren, 2004; Kelkar et al., 2008). It is well established that microsatellites with a larger number of repeats have higher mutation rate due to increased probability of slippage (Ellegren, 2004). In addition, other intrinsic features of a microsatellites such as the repeated motif (andits composition) can affect its mutation rate. Different levels of genetic variability depending on microsatellite motif have been reported in some eukaryote genomes, and the relative mutation rate of dinucleotides versus tri- and hepta-nucleotides is estimated to be between 1.5 and 2.1 (Anderson et al., 2000; Chakraborty et al., 1997). This is thought to be due to both the differential mutation rate during replication and to a higher rate of recombination and consequent mismatch repair

(Chakraborty et al., 1997; Li et al., 2002). It was observed that longer microsatellites showed a major polymorphism as well as loci with shorter repeat motif had higher average repeat numbers (Chakraborty et al., 1997; Ellegren, 2004; Kelkar et al., 2008).). In this study, di-nucleotide microsatellites showed an higher number of repeats and thus they had higher mutation rate than tri-nucleotides or hepta-nucleotides for both A-SSRs and E-SSRs. This observation was in agreement with a previous study in *Posidonia oceanica* in which it has been showed that the relative mutation rate of di-nucleotides versus tri-nucleotides was 2.5 to 3 times higher (Arnaud-Haond et al., 2005). However in the present study the comparison between the two types of markers showed that E-SSRs present a restricted repeat array in comparison to that showed for genomic microsatellites for the same microsatellites classes.

As consequence of their major number of alleles and higher allele frequencies, A-SSRs revealed higher discrimination power (D = 0.613 ± 0.318) than E-SSRs (D =0.394 ± 0.294). Although both markers sets showed a moderate informativeness level (0.25 < PIC < 0.5), A-SSRs have higher polymorphic information (PIC = 0.481 ± 0.265) than E-SSRs (PIC = 0.318 ± 0.219), Consequently, A-SSRs detected a higher clonal diversity in the populations analyzed than E-SSRs.

According to D-classes and PIC-classes until the 11th position, there was an almost precise correspondence in the order of microsatellites; 7 A-SSRs di-nucleotides (Alberto et al., 2003) and 4 E-SSRs (two di-nucleotides and two-tri-nucleotides) are the most discriminating microsatellites. Interestingly, the same 7 A-SSRs di-nucleotides were identified as the optimal combination of markers to assess clonality in *Posidonia* populations in a previous study (Arnaud-Haond et al., 2005).

Also, it has been showed that the use of di-nucleotides allowed a higher estimate of the clonal diversity in comparison to the use of the same number of tri-nucleotide and hepta-nucleotide microsatellites on the same population (Arnaud-Haond et al., 2005).

This relationship between the discrimination power and the repeat motif was observed also in this study for A-SSRs, where, following the D-classes, all di-nucleotides appeared in the higher classes followed by all tri-nucleotides in the intermediate positions and by the epta-nucleotide at last position. For the E-SSRs, a different pattern emerged because di-nucleotides and tri-nucleotides alternated along the D-classes not showing a correspondence between the repeat motif and the discrimination power. Probably, functional constraints in the EST sequences from where E-SSRs come, affected the microsatellites mutability in a way that is independent from the repeat motif.

In general, the comparison between A-SSRs and E-SSRs showed that different types of markers lead to different conclusions on the level of genotypic richness. Considering the A-SSR set, for example, Marettimo emerged as one of the most genetically diverse population, while a different result was obtained with the E-SSRs, where Meloria appeared as the population with the highest number of distinct individuals.

Although differences are not striking, the results confirm that A-SSR markers are more variable than putatively functional markers and consequently are more powerful than E-SSRs in the assessment of clonal diversity. However, although the A-SSRs detect a higher genotypic variation than the E-SSRs, the combined set of markers allowed the best discrimination of the genets than either independent set

of markers alone (Appendix). Considering all the six populations together, the saturation in the discrimination of individuals is reached after 27 loci in the less polymorphic population of Lacco Ameno (R = 0.808) while a lower number of primers are sufficient for the saturation in more polymorphic populations. For example, 4 loci are sufficient to saturate the clonal diversity in Marettimo, 7 loci in Delimara and 14 in Meloria. Thus the use of all 27 SSRs could represent a good tool to assess accurately the clonality in *Posidonia oceanica* populations also different from those used in this study.

A-SSRs also showed a higher resolution power in the analysis of gene flow and population genetic structure. Pair-wise F_{ST} estimates from A-SSRs clearly suggested the presence of differentiation that followed a trend related to the geographical position of the populations with southern population more differentiated from northern population. This pattern was not well revealed with the analysis of E-SSRs. The PCoA analysis confirmed that A-SSRs accounted for a major percentage of the total variation in comparison to E-SSRs.

In addition, EST-derived SSRs and anonymous SSRs generated very different results in estimating the number of distinct populations represented in the total pool of individuals genotyped. A-SSRs appeared the most powerful markers identifying 15 clusters vs. 9 cluster identified with E-SSRs. At level of single populations, it was interesting to note the most of the populations showed a clear distinction between samples collected at the two depths. This pattern was better highlighted by A-SSRs, except for Stareso, where the distinction between depths was revealed by E-SSRs.

3.6-Conclusions

Population genetic parameters generated by A-SSRs and E-SSRs are very different. A lower level of allelic diversity was found in E-SSRs compared with genomic A-SSRs. Overall, A-SSRs also had a major discrimination power in assessment of clonality in *Posidonia oceanica*.

For the analysis of population differentiation (F-statistics) the A-SSRs have much higher power than E-SSRS and alone are efficient to have reliable estimate of these measures. A-SSRs performed better than E-SSR in population cluster analysis in comparison with E-SSRs. This is also true at within-population level. On the other hand, despite showing lower levels of polymorphisms, the EST-derived SSRs, reflecting the genetic diversity inside or adjacent to the genes (Varshney et al., 2005), have much more potential than genomic SSRs to reveal the functional variation between individuals. This characteristics makes the E-SSRs specifically used for studying selection processes in fact they may be under selective pressure or subject to background selection caused by genetic hitchhiking.

The lower level of diversity detected at these markers in this study could just suggest that they may not behave as effective neutral markers and that selection could have affected the performance of the microsatellites in this population genetic analysis.

In the next chapter, the whole set of markers will be utilized to search for loci under selection (outliers) along the bathymetric and latitudinal gradient.

3.7-References

Agarwal, M., Shrivastava, N., Padh, H. (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Rep,* 27: 617–631.

Alberto, F., Correia, L., Arnaud-Haond, S., Billot, C., Duarte, C. M., & Serrao, E. (2003). New microsatellite markers for the endemic Mediterranean seagrass *Posidonia oceanica. Molecular Ecology Notes,* 3(2): 253–255.

Anderson, T. J. C., Su, X. Z., Roddam, A., Day, K. P. (2000). Complex mutations in a high proportion of microsatellite loci from the protozoan parasite *Plasmodium falciparum. Mol Ecol,* 9: 1599–1608.

Arnaud-Haond, S., Alberto, F., Teixeira, S., Procaccini, G., Serrao, E.A. & Duarte, C.M. (2005). Assessing genetic diversity in clonal organisms: low diversity or low resolution? Combining power and cost-efficiency in selecting markers. *Journal of Heredity,* 96: 434–440.

Arnold, C., Rossetto, M., McNally, J., Henry, R. J. (2002). The application of SSRs characterized for grape (*Vitis vinifera*) to conservation studies in Vitaceae. *American Journal of Botany,* 89: 22–28.

Botstein, D., White, R. L., Skolnick, M. and Davis, R. W. (1980). Construction of genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet,* 32: 314-331.

Buonaccorsi, V. P., Kimbrell, C. A, Lynn, E. A., Hyde, J. R. (2012). Comparative population genetic analysis of bocaccio rockfish *Sebastes paucispinis* using anonymous and gene-associated simple sequence repeat loci. *The Journal of heredity,* 103(3): 391–9.

Chakraborty, R., Kimmel, M., Stivers, D. N., Davison, L. J., Deka R. (1997). Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proc Natl Acad Sci USA,* 94: 1041–1046.

Corander, J. and P. Marttinen. (2006). Bayesian identification of admixture events using multilocus molecular markers. *Mol Ecol.,*15(10): 2833-43.

Coulibaly, I., Gharbi, K., Danzmann, R. G., Yao, J., Rexroad, C. E. (2005). Characterization and comparison of microsatellites derived from repeat-enriched libraries and expressed sequence tags. *Animal genetics,* 36(4): 309–15.

Dieringer, D and C. Schlötterer. (2003). Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes,* 3: 167-169.

Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. *Nat Rev Genet.,* 5: 435–445.

Eujayl, I., Sorrells, M., Baum, M., Wolters, P., Powell, W. (2001). Assessment of genotypic variation among cultivated durum wheat based on EST-SSRs and genomic SSRs. *Euphytica,* 119: 39–43.

Eujayl, I., Sledge, M. K., Wang, L., May, G. D., Chekhovskiy, K., Zwonitzer, J. C., Mian, M. A. R. (2004). Medicago truncatula Est-SSRs reveal cross-species genetic markers for medicago spp. TAG Theoretical and Applied Genetics, 108 (3): 414-422.

Finger, A. and C. Klank. (2010). Molecular methods - Blessing or curse? In: J.C. Habel & T. Assmann (Eds) Relict Species - *Phylogeography and Conservation Biology*. Springer Heidelberg, p. 309-320.

Gao, L. F, Jing, R. L, Huo, N. X, Li, Y., Li, X. P., Zhou, R. H., Chang, X. P., Tang, J. F., Ma, Z. Y., Jia, J. Z. (2004). One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat. *Theor Appl Genet.*, 108: 1392–1400.

Goldstein, D. and C. Schlötterer. (1999). Microsatellites: Evolution and applications. Oxford University Press, Oxford, UK.

Goudet, J. (2001). FSTAT, A Program to Estimate and Test Gene Diversities pool Daphnia metapopulations. Ecography 26: 731–738. and Fixation Indices, Version 2.9.3. http://www.unil.ch/izea/soft Pamilo, P., S. Palsson and O. Savolainen, 1999 Deleterious mutawares/fstat.html.

Hu, J., Wang, L., Li, J. (2011). Comparison of genomic SSR and EST-SSR markers for estimating genetic diversity in cucumber. *Biologia Plantarum*, 55(3): 577–580.

Kalia, R. K, Rai, M. K., Kalia, S., Singh, R., Dhawan, A. K. (2011). Microsatellite markers: an overview of the recent progress in plants. *Euphytica*, 177: 309-334

Kelkar, Y. D., Tyekucheva, S., Chiaromonte, F., Makova, K. D. (2008). The genome-wide determinants of human and chimpanzee microsatellite evolution. *Genome Res.*, 18: 30–38.

Kim, K. S., Ratcliffe, S. T., French, B. W., Liu, L., Sappington, T. W. (2008). Utility of EST-derived SSRs as population genetics markers in a beetle. *The Journal of heredity*, 99(2): 112–24.

Lehmann, T., Hawley, W. A., Collins, F. H. (1996). An evaluation of evolutionary constraints on microsatellite loci using null alleles. *Genetics*, 144: 1155–1163.

Li, Y.C., Korol, A. B., Fahima, T., Beiles, A., Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol Ecol.*, 11: 2453–2465.

Metzgar, D., Bytof, J., Wills, C. (2000). Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Res.*,10: 72–80.

Migliaccio, M., De Martino, F., Silvestre, F., Procaccini, G. (2005). Meadow-scale genetic structure in *Posidonia oceanica*. *Marine Ecology Progress Series*, 304: 55–65.

Molina-Luzón, M. J., López, J. R., Navajas-Pérez, R., Robles, F., Ruiz-Rejón, C., De La Herrán, R. (2012). Validation and comparison of microsatellite markers derived from Senegalese sole (Solea senegalensis, Kaup) genomic and expressed sequence tags libraries. *Molecular ecology resources*, 12(5): 956–66. doi:10.1111/j.1755-0998.2012.03163.x

Oliveira, E. J.; Pádua, J. G.; Zucchi, M. I.; Vencovsky, R. And Vieira, M. L. C. (2006). Origin, evolution and genome distribution of microsatellites. *Genetics and Molecular Biology*, 29 (2): 294-307.

Park, S. D. E. (2001). Trypanotolerance in West African Cattle and the Population Genetic Effects of Selection [Ph.D. thesis] University of Dublin

Peakall, R., Smouse, P. E. (2006). GENEALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6: 288-295.

Pemberton, J. M., Slate, J., Bancroft, D. R., Barrett, J. A. (1995). Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology*, 4: 249–252.

Procaccini, G. and M. Waycott. (1998.) Microsatellite loci identified in the seagrass *Posidonia oceanica* (L.) Delile. *Journal of Heredity*, 89: 562–568

Scott, K. D., Eggler, P., Seaton, G., Rosetto, M., Ablett, E. M., Lee, L. S., Henry, R. J. (2000). Analysis of SSRs derived from grape ESTs. *Theor Appl Genet*, 100: 723–726

Selkoe, K. and R. Toonen. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, 9: 615-629.

Simko, I. (2009). Development of EST-SSR markers for the study of population structure in lettuce (*Lactuca sativa* L.). *The Journal of heredity*, 100(2): 256–62.

Song, Y., Jiang, X., Zhang, M., Wang, Z., Bo, W., An, X., Zhang, D., Zhang, Z.(2012). Differences of EST-SSR and genomic-SSR markers in assessing genetic diversity in poplar. *Forestry Studies in China*, 14(1): 1–7.

Squirrel, J., P.M. Hollingsworth, M.L. Woodhead, J. Russell, A.J. Lowe, M. Gibby, and W. Powell. (2003). How much effort is required to isolate nuclear microsatellites from plants? *Molecular Ecology*, 12: 1339-1348..

Tang, J., Hanage, W. P., Fraser, C., Corander, J. (2009). Identifying currents in the gene pool for bacterial populations using an integrative approach. *PLoS Comput Biol*, 5:e1000455.

Tautz, D. (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, 17: 6463–6471.

Tessier, C., David, J., This, P., Boursiquot, J. M., Charrier, A. (1999). Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. *TAG Theoretical and Applied Genetics*, 98(1): 171–177. doi:10.1007/s001220051054

Thiel, T., Michalek, W., Varshney, R. K., Graner, A. (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *TAG. Theoretical and applied genetics*, 106(3): 411–22.

Varshney, R. K., Graner, A., Sorrells, M. E. (2005). Genic microsatellite markers in plants: features and applications. *Trends in biotechnology*, 23(1): 48–55.

Vigouroux, Y., McMullen, M., Hittinger, C. T., Houchins, K., Schulz, L., Kresovich, S., Matsuoka, Y., Doebley, J. (2002). Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication. *PNAS*, 99: 9650-9655.

Vilas, R., Bouza, C., Vera, M., Millán, A, Martínez, P. (2010). Variation in anonymous and EST-microsatellites suggests adaptive population divergence in turbot. *Marine Ecology Progress Series*, 420: 231–239.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Freijters, A., Pot, J., Peleman, J., Kuiper, M. & Zabeau, M. (1995). AFLP: a new concept for DNA fingerprinting. *Nucleic Acids Res*, 21: 4407-4414.

Wiehe, T. 1998. The effect of selective sweeps on the variance of the allele distribution of a linked multiallele locus: hitchhiking of microsatellites. Theor. Popul. Biol. 53:272–283.

Williams, J. G, Kubelik, A. R., Livak, K. J., Rafalski, J. A., Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, 18: 6531–6535.

Woodhead, M., Russell, J., Squirrell, J., Hollingsworth, P. M., Mackenzie, K., Gibby, M., Powell, W. (2005). Comparative analysis of population genetic structure in *Athyrium distentifolium* (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. *Molecular ecology*, 14(6): 1681–95.

Zietkiewicz, E., Rafalski, A., Labuda, D. (1994). Genome fingerprinting by simple sequence repeats (SSR)-anchored PCR amplification. *Genomics*, 20: 176–183.

3.8-Appendix

The performance of the combined set of markers in the assessment of clonal diversity was assessed with the aim to identify the most efficient marker set to use in future estimates of clonality in *P. oceanica*. In comparison to the A-SSR set and the E-SSR set used separately, the combined set of the two classes of markers increased the overall value of genotypic richness in the global set analyzed discriminating 251 individuals on 271 multilocus genotypes (R = 0.926, Table 3.A1). Much higher clonal diversity was revealed using the combined set of markers: the levels of diversity ranged from 0.808 (Lacco Ameno) to 1 (Delimara and Meloria; mean R across populations = 0.933 ± 0.073) (Table 3.A1 and Figure 3.A1).

In order to select the best microsatellite combination to saturate genotypic diversity, R values obtained adding one locus per time, starting from the one with highest D values to the lowest one (see Table 3.6 in the main text), were plotted (Figure 3.A2).

R values saturate after 27 loci in the less polymorphic population of Lacco Ameno, while it saturate with 4, 7 and 14 loci in Marettimo, Delimara and Meloria, respectively. Considering al populations together, saturation is reached after 27 loci (Figure 3.A2).

Table 3.A1. Clonal diversity in the *Posidonia oceanica* populations studied. G = number of distinct multilocus genotypes; N= number of individuals; R (G-1/N1) = genotypic richness. Overall = analysis including the entire data set; SD = standard deviation. For comparative purpose also the separate analysis for the A-SSR and E-SSRs, already reported in the Chapter, is presented.

	A-SSR set			E-SSR set			Combined set		
	G	N	R	G	N	R	G	N	R
Overall	217	271	0.800	182	271	0.670	251	271	0.926
Delimara	39	40	0.974	26	340	0.641	40	40	1.000
Lacco Ameno	33	53	0.615	32	53	0.596	43	53	0.808
Marettimo	43	44	0.977	28	44	0.628	43	44	0.977
Meloria	38	40	0.949	31	40	0.769	40	40	1.000
Piombino	37	43	0.857	32	43	0.738	41	43	0.952
Stareso	27	51	0.520	33	51	0.640	44	51	0.860
Mean			0.815			0.669			0.933
SD			0.182			0.063			0.073

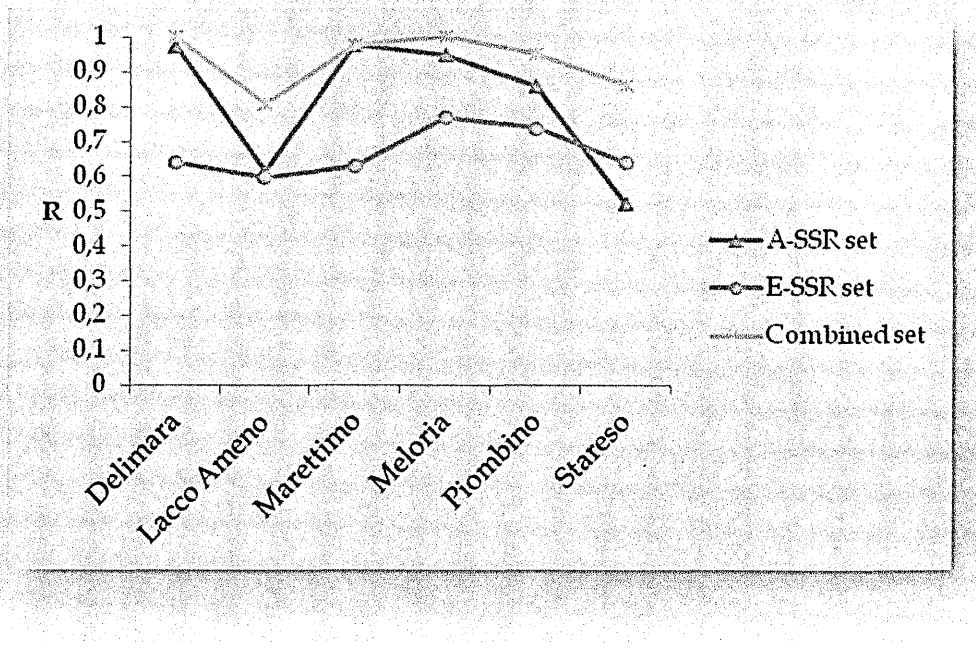


Figure 3.A1. Genotypic richness at population level for the A-SSRs, E-SSRs and the combined set of microsatellites. R = genotypic richness.

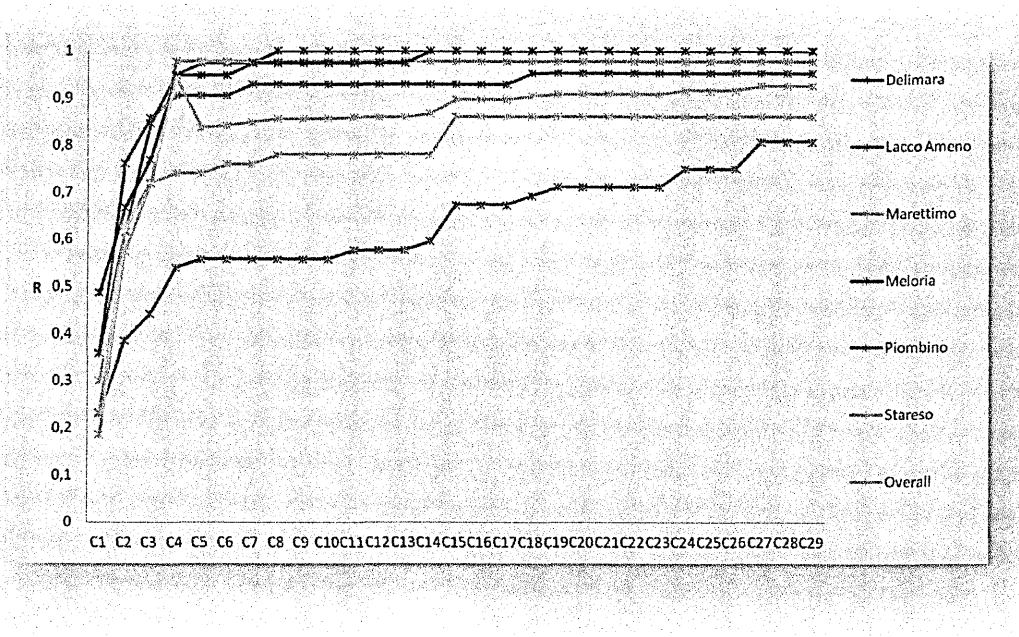
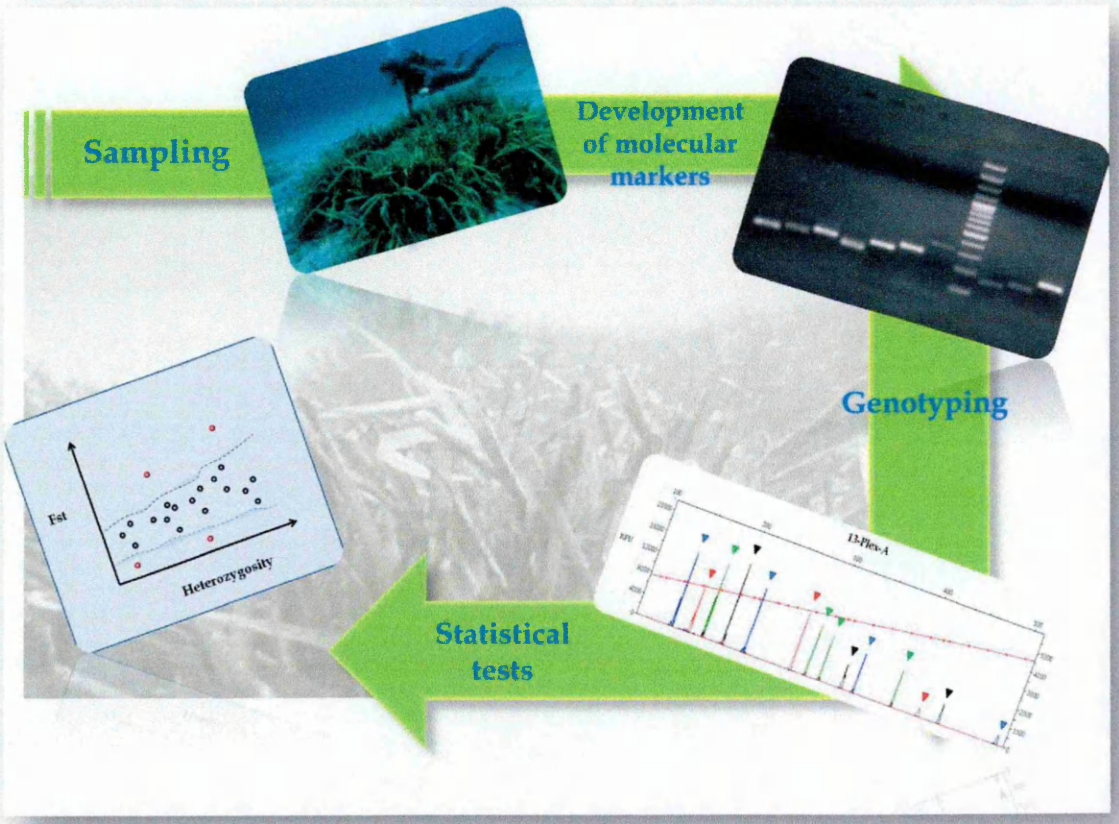


Figure 3.A2. Values of genotypic richness (R) per population and overall, obtained adding sequentially the SSR loci with most discriminating to the less discriminating power, following the D-classes reported in Table 3.7.

Chapter 4


Sampling
Development of molecular markers
Genotyping
Statistical tests
Fst
Heterozygosity

4-Genome scan along depth and latitudinal gradients in the seagrass *Posidonia oceanica*

4.1-Abstract

Natural landscapes are a mixture of biotic and abiotic challenges that result in a complex mosaic of spatially-varied selective pressures. These pressures are often exacerbated by anthropogenic activities that expose natural populations to novel selective pressures and with unusual intensity. Natural populations respond to these challenges, including climate change, shifting their occupational range,adapting to the changing environment or with local extinction. Resilient species like plants are the best candidates to study adaptive responses to severe changes in the environment given their sessil habit. Natural plant populations can adapt to the changing environment through phenotypic plasticity or genetic adaptation. Here, signature of selection underlying adaptive responses of natural *P. oceanica* populations to light and temperature changes was investigated. To identify such signatures a genome scan approach comparing populations along bathymetric and latitudinal gradients was used. Two putative candidate genes involved in the photosynthetic pathway and in the process of translation, and a chloroplast intron were identified. These findings allowed us to gain insights into the genetic adaptation of *P. oceanica* to light and temperature changes and will be of relevance in the design of conservation plans for this species.

Keywords: light, temperature, global change, local adaptation, selection, microsatellites, genetic hitch-hiking, outlier loci, *Posidonia oceanica*, genome scan.

4.2-Introduction

Species across the globe are experiencing drastic changes in environmental conditions as a result of human activities (Hansen, 2012). For seagrasses, the two main human impacts are represented by sediment loading and eutrophication that have major effects on water clarity, by reducing the light available for photosynthesis (Orth et al., 2006, Duarte et al., 2008, Waycott et al., 2009). Threats for marine plants are also represented by climate change-related increase in sea temperature (Orth et al. 2006). Temperature stress on seagrasses resulted in distribution shifts, changes in patterns of sexual reproduction, altered growth rates, metabolism, and carbon balance (Short et al., 2001, Short and Neckles, 1999). Elevated temperatures may also increase the growth of competitive algae and epiphytes, which can overgrow seagrasses and reduce the available light they need to survive (Björk et al., 2008). With future global warming, there may be a 1-5 m rise in the seawater levels by 2100 (Overpeck et al., 2006, Hansen, 2007, Rahmstorf, 2007). Rising sea levels may impact seagrass communities due to increases in water depths above present meadows (thereby reducing light), changing currents and causing erosion and increased turbidity in rivers (Short et al., 2001).

The increase in temperature, sediment loading and eutrophication are not the only threats for seagrasses. Climate change also increases the intensity (IPCC 2007) and frequency (Trenberth, 2005) of tropical storms with consequent increase in turbidity that can remain long after the storm subsides, causing shading-out of

plants. These altered weather patterns may cause more clouds to form, thus reducing photosynthetically active radiation (PAR) of visible light and representing a stressespecially for those seagrass meadows that already grow close to their low-light limit (Björk et al., 2008). In addition, the increase in oceanic CO_2, as consequence of the increase in atmospheric CO_2 concentration (Harley et al. 2006), affects seagrass photosynthesis and growth. Higher CO_2 levels may boost the production and biomass of epiphytic algae on seagrass leaves, which may adversely impact seagrasses by causing shading (Beer and Koch, 1996).

Thus, although multiple stresses or threats act on the seagrass ecosystems, most of them directly or indirectly affect the light availability for photosynthesis. In addition to the changes in light, the increase in water temperature has not to be underestimated, both representing limiting factors for the growth of seagrasses. Light drives photosynthesis and temperature affects both photosynthesis and respiration and together with light regulate seagrasses seasonal growth. Light availability and temperature affect seagrasses across a range of spatial scales, including individual leaf responses, shoot-scale responses and alterations to the meadow structure (Olesen et al., 2002).

In the critical scenario of climate change it is important to understand how seagrasses respond to fluctuations of these two factors in order to establish appropriate measurement of management and conservation. Organisms response to global environmental change can be ecological or evolutionary. The former includes phenotypic plasticity and dispersal, while the latter involves genetic change (Reusch and Wood, 2007). In the context of the evolutionary response to global change, it is evident that environmental change at all spatial scales is

rapidly altering selection regimes for global flora and fauna (Reusch and Wood, 2007). Selective pressures lead local populations to have resident genotypes with higher fitness, in respect to local environmental conditions, than genotypes originating from different populations (Kawecki and Ebert, 2004). This is the process termed local adaptation that results from the balance between gene flow and many natural selection factors, including climate (Savolainen et al. 2007; Hoffmann and Willi, 2008). Elucidating the genetic basis of local adaptive evolution of natural populations by identifying loci under local selection may become crucial for understanding population fate and for managing biodiversity using the adaptive potential of species.

Posidonia oceanica meadows offer an good opportunity to gain insights into habitat dependent selection, since meadows can occur along a wide bathymetric gradient (from the surface to 45 m depth, in very clear waters), and at latitudes encompassing the whole Mediterranean basin. This implies that the plants living at different depths and different latitudes experience different environmental conditions, mainly as regards as light and temperature, two limiting factors to the growth of this plants. For photosynthesis to occur in seagrasses, light must penetrate the water column, enter the canopy of leaf blades, pass through a layer of epiphytes on the surface of the leaf and finally enter the leaf epidermis to reach the photosynthetic apparatus (Dalla Via et al., 1998). Light is attenuated at each of these steps. Also, light intensity rapidly decrease with water depth by which plants in deep have to cope with higher attenuation of the light than plants living in shallow waters (Ralph et al., 2007). Because of that, seagrasses are particularly sensitive to reduction in light availability whereby a small decreases of light can

cause significant decline in growth and distribution, especially at the deepest edge of the meadow where plants are at the limit of their minimum light requirements. In addition to the attenuation of the light with the depth, the increase in water turbidity resulting from human activities and/or or by indirect effect of climate change exacerbates the consequences for the *P. oceanica* meadows.

The objective of this study is the identification of loci under selection related to depth and latitude that reflect changes in light and temperature regimes. This is provided by the means of a genome scan employing microsatellites markers. Two classes of microsatellite markers were utilized: genomic and EST-linked. Genomic markers are not mapped in the genome, while EST-linked microsatellites are flanking to a gene of known function (see previous chapters). Hence, EST-linked markers represent a very useful means to identify functional genes possibly under selection (Vasemagi et al., 2005; Namroud et al., 2008).

Genome scans consist in the screening of genome-wide patterns of DNA polymorphisms and in the application of statistical tests to identify loci under selection by allowing to distinguish between neutral loci, loci that are not evolving directly in response to selection, the dynamics of which are controlled mainly by genetic drift and migration, from outlier loci, loci that show behavior or pattern of variation that are extremely divergent from the rest of genome (locus-specific effects; Luikart et al., 2003, Storz, 2005). Genome scans are based on the concept of genetic hitch-hiking, by which if molecular markers are physically linked to functionally important genes, the action of selection acting on such genes also affects the flanking markers (Maynard and Haigh, 1974). In non-model genetic species, with limited or no genome information, genome scans are a promising

approach to detect local adaptive genetic variation, since it has the advantage of not requiring knowledge of genetic divergence in quantitative traits between populations. It also does not require a complete knowledge of DNA sequences underlying the loci investigated (Storz, 2005).

Actually, the application of the genome scan is a means to have a major knowledge about the fitness variation in natural populations and to establish causal links between genotype, phenotype, and fitness in natural populations (Dalziel et al., 2009). Once selected genes are identified, it is necessary to work up through the levels of biological organization to fully understand the consequences of genetic variation at these candidate genes. The genetic variation at the selected loci, in fact, can affect the function and /or amount of proteins, which may then alter protein–protein interactions, influence biochemical pathways and networks and eventually modify cellular function, organismal phenotype, whole-organism performance capabilities and fitness (Dalziel et al., 2009).

In plants, adaptation to environmental gradients (Bonin et al., 2006; Jump et al., 2006, Wang et al., 2012) and adaptation to contrasting habitats (Mealor and Hild, 2006; Oetjen and Reusch, 2007; Oetjen et al., 2009; Namroud et al., 2008, Meyer et al., 2009), have been addressed by means of genome scans. All these studies reported evidence for divergent selection on at least some marker loci. The proportion of adaptive loci, which exhibit differentiation stronger than expected given the genome-wide differentiation, was generally found between 2% and 15% (Meyer et al., 2009) of the total number of loci utilized. Among other species, a genome scan approach has been applied to the seagrass species *Zostera marina*. Oetjen and Reusch (2007) used 25 microsatellites markers to search for signals of

positive selection between populations living in different habitats and identified 3 outlier loci. Subsequently, the same authors (Oetjen et al., 2009) performed a genome scan increasing the number of microsatellites and introducing SNPs, a new type of marker for seagrasses. Starting from 46 markers (6 SNPs, 16 EST-microsatellites and 11 anonymous microsatellites) the authors identified 6 outlier loci (Oetjen et al., 2009).

Here the hypothesis is that different environmental conditions could act as selective forces, leading to local adaptation. Because *P. oceanica* meadows extend along a latitudinal gradient in the Mediterranean sea, our sampling was designed to compare not only meadow's stands sampled at different depths but also meadows sampled at different latitudes, where temperature and light availability can also change significantly.

Hence, we aim to test the following specific hypotheses:

1. The depth-related genetic structure previously observed in Lacco Ameno (Migliaccio et al., 2005) has leaded to local adaptation.
2. A small proportion of the loci utilized will be under selection. Since we considered two environmental gradients (light and temperature) outlier loci should be involved in the photosynthesis or in the response to external stimuli.
3. The same outlier loci will be found along the two environmental gradients considered (depth and latitude).

We also have to consider that the populations analyzed are sampled in natural condition and additional factors, different from light and temperature, could

underly the adaptive variation we analyze and could affect the results of the genome scan along the two environmental gradients.

4.3-Materials and Methods

4.3.1-Sampling design and DNA extraction

P. oceanica samples were collected along a latitudinal (Figure 4.1) and a bathymetric gradient. For each geographic location populations were sampled at two different depths (5 m and 25 m), above and below the summer thermocline (about 15 m depth). Between twenty and thirty samples, consisting of adult leaves, were sampled at each depth, for each of the six localities (N, number of samples = 40 to 53, shown in Figure 4.1). Leaves were cleaned from epiphytes and stored in silica gel prior DNA extraction. Genomic DNA was extracted using the NucleoSpin® 96 Plant II kit following manufacturer instructions (Macherey-Nagel, see also chapter 2).

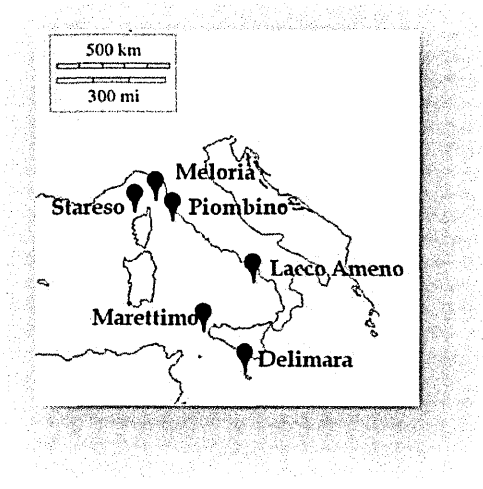


Figure 4.1. Geographical position of the 6 *Posidonia oceanica* populations sampled. Delimara - Strait of Sicily, Malta (35° 51' 28" N; 14° 33' 07" E, N = 40); Lacco Ameno - Central Tyrrhenian Sea, Italy (40° 45' 52'' N; 13° 53' 29'' E, N = 53); Marettimo - Southern Tyrrhenian Sea, Italy (37° 58' 10" N; 12° 04' 76"E, N = 44); Meloria - Tyrrhenian Sea, Italy (43° 31' 41" N; 10° 10' 32" E, N = 40). Piombino - North Tyrrhenian Sea, Italy (42° 56' 92"N; 10° 38' 99"E, N = 43); Stareso -Corsica Sea (8° 45'E; 42° 35' N, N = 51).

4-3.2-Outlier detection: criteria and methods

In total, 29 microsatellites were used in this study, including 13 anonymous (Procaccini and Waycott, 1998; Alberto et al., 2003) and 16-EST-linked microsatellites (see Chapter 2).

In order to identify outlier loci linked to temperature and light gradients, (sub)populations from stands located at different depths in the same geographic location and populations from different latitudes were compared.

To identify outlier loci, three methods were used: Beaumont and Nichols (1996) Fdist approach, implemented in the software LOSITAN (Antao et al., 2008), the hierarchical Bayesian method BAYESCAN (Foll and Gaggiotti, 2008) and the Schlötterer tests (Schlötterer, 2002a, b; Schlötterer and Dierienger, 2005, Kauer et al. 2003).

LOSITAN simulates a null distribution of F_{ST} values based on an infinite island model of populations. Loci with an F_{ST} deviating from neutral expectations are considered outlier loci. In this study, the neutral distribution of F_{ST} was simulated with 1,000,000 iterations and the two options "neutral mean F_{ST}" and "forced mean F_{ST}" were chosen following Antao et al. (2008).

BAYESCAN directly infers the posterior probability of each locus to be under the effect of selection by comparing two alternative models, which alternatively include or exclude the effect of selection (Foll and Gaggiotti, 2008). The resulting posterior probabilities (p) are used to define posterior odds (PO), that determine if the model including or excluding selection has the highest likelihood (Foll and Gaggiotti, 2008). Jeffreys scale (1961) is then used to define the threshold at which the model is accepted: PO > 3 substantial ($\log_{10}$PO>0.5); PO > 10 strong

(log_{10}PO>1.0); PO > 32 very strong (log_{10}PO>1.5) and PO >100 decisive evidence for selection. In the present analysis a threshold of PO >10 (strong) was used for a marker to be considered under selection, that corresponds to a posterior probability greater than 0.91 for the model accounting for selection. BAYESCAN was run with the following settings: 20 pilot runs, length of pilot runs 5000 resulting in 100,000 iterations, burn-in of 50,000 iterations.

Schlötterer tests (Schlötterer, 2002a, b; Schlötterer and Dierienger, 2005; Kauer et al. 2003) are an alternative multilocus outlier tests that uses summary statistics other than F_{ST}. Using the software Microsatellite analyzer (MSA) ver. 4.05 (Dieringer and Schlötterer, 2003), the variance in repeat number, heterozygosity and θ values for each locus in each population pairs were calculated. Using EXCEL, the standardized measures for the statistics (lnRH, lnRV and lnRθ) were estimated. It has been demonstrated that these statistics are approximately normally distributed under neutrality (Kauer, et al., 2003) and thus, after standardization (mean = 0; SD = 1), 95% of neutral loci are expected to have values between -1.96 and 1.96 (99% between -2.58 and 2.58; 99.9% between -3.29 and 3.29). In the cases when a locus was monomorphic in one of the populations utilized in the pair-wise comparison, a single different allele was added to the sample in order to avoid heterozygosis, variance in the number of repeat and θ value being equal to zero.

Outlier loci were identified based on the following criteria:

-To reduce type I error, only loci that were detected in several pair-wise analysis of population comparisons were considered as real outliers. Also, the outlier detection based on population differentiation is sensitive to signatures of

population history (Luikart et al., 2003). One way to deal with these confounding factors includes the identifications of population clusters (Holderegger et al., 2008). The cluster analysis conducted on 12 anonymous microsatellites using BAPS ver. 6.0 (Tang et al., 2009) and described in chapter 3, identified 15 population clusters, corresponding largely to the 6 superficial and deep stands sampled in correspondence of each population and excluding the population of Meloria that did not show a clear bathymetric structure. In total, six pair-wise population comparisons were performed including 5 pair-wise comparisons for each population showing a clear bathymetric structure and a global analysis contrasting pooled shallow and deep stands (Table 4.1). For the analysis of outlier loci along the latitudinal gradient pair-wise comparisons involving the same population were excluded. Three population pairs were performed selecting the pairs of populations, following the criteria that in each comparison a population from the Southern (Figure 4.1 and Table 4.1) and one from the Northern basin were contrasted and that each population was compared only once. In addition the pooled southern and northern populations were contrasted.

-To decide on the minimum threshold needed to define an outlier as a locus potentially under selection, all loci were plotted in a frequency distribution table, based on the frequency with which they were detected as outliers in the pair-wise population comparisons. Loci falling outside the 90% boundaries of this frequency distribution were considered as real outliers and used for further analyses.

-Only loci repeatedly detected across three runs of the same pair-wise population comparison were considered for further analysis.

4.3.3-Population structure and differentiation at neutral and selected loci

F_{ST} values from neutral and selected loci were calculated using the software FSTAT ver 2.9.3 (Goudet, 2001).

To assess the component of genetic variance due to habitat differences, an analysis of molecular variance (AMOVA) was carried using Arlequin ver. 3.5 (Excoffier et al., 2005). The AMOVA was performed separately for the neutral and selected loci classified a posteriori according to the outlier tests and three hierarchical levels (among groups, among populations within groups and within populations) were considered.

To visualize the spatial distribution of the populations a PCoA (Principal Coordinate Analysis) with neutral and selected loci was performed using GenAlEx ver. 6.5 (Peakall and Smouse, 2006).

4.4-Results

4.4.1-Outliers analysis

The genome scan performed along the bathymetric and latitudinal gradient generated non overlapping results.

Following the criteria outlined in Material and Methods and contrasting populations along the depth gradient no outliers were identified (Table 4.1).

The outlier analysis conducted along the latitudinal gradient and involving independent population pairs (see Material and Methods) identified 3 outlier loci (highlighted in green in Table 4.1).

A summary of the outliers analysis performed along both gradients with three different approaches (LOSITAN, BAYESCAN and Schlötterer tests) is reported in the appendix. Along the depth gradient the three outlier tests were not concordant in the identification of the same outlier loci (Table 4 A.3). Within a single outlier test, even if the same locus was identified in more pair-wise comparisons (i.e. three out six comparisons) its distribution frequency was not high enough to statistically support its identification as outlier . Conversely, along the latitudinal gradient the three outlier tests not only showed a convergence in the identification of the same three outlier loci (Poc-*trn*, Pooc-153 and Pooc-PC045G11, Table 4 A.4) but also they identified these loci as outlier in almost all the pair-wise comparisons examined (three out four comparisons).

Table 4.1. Outlier inference along the latitude and depth gradient. 1 = Indicates loci detected as outliers in the specific pair-wise comparison. Sha = shallow (sub)population; Dee = deep(sub)population; frequency = number of times a outlier locus is detected in a particular comparison. For the latitude gradient in green are outlier loci for latitude (>2 comparisons).

	Depth gradient							Latitude gradient				
Pair-wise comparisons	**Delimara Sha -Dee**	**Lacco Sha-Dee**	**Marettimo Sha-Dee**	**Piombino Sha-Dee**	**Stareso Sha-Dee**	**Pool Sha-Dee**	**Frequency**	**Delimara-Piombino**	**Lacco-Meloria**	**Marettimo-Stareso**	**Southern-Northern**	**Frequency**
Locus												
Poc-45	1		1	1			3					0
Po-5							0					0
Poc-5							0	1				1
Poc-35	1			1	1		3	1				1
Po5-49					1		1					0
Poc-trn							0		1	1	1	3
Poc-26						1	1				1	1
Po15							0					0
Po5-10	1				1		2					0
Po4-3							0					0
Po5-39				1	1		2			1		1
Po5-40							0				1	1
Poc-42	1			1			2		1			1
Pooc-229					1		1					0
Pooc-264		1		1		1	3	1	1			2
Pooc-330							0		1	1		2
Pooc-3	1						1		1			1
Pooc-54	1	1					2	1			1	2
Pooc-153		1	1			1	3		1	1	1	3
Pooc-333							0				1	1
Pooc-125			1			1	2		1			1
Pooc-PC045G11					1		1	1	1		1	3
Pooc-PC044B02				1			1					0
Pooc-214		1					1					0
Pooc-300							0	1				1
Pooc-PC047G07			1	1			2					0
Pooc-50	1		1	1			3					0
Pooc-361	1						1					0
Pooc-PC003H09							0		1			1

4.4.2-Genetic variation at selected and neutral loci

Genetic variation at neutral loci and loci putatively under selection was analyzed. Loci under selection were the outlier loci identified following the criteria listed in Material and Methods.

Neutral Pair-wise F_{ST} estimates between the shallow and deep site at each population are presented in Table 4.2 (highlighted in yellow). All pair-wise F_{ST} values indicated highly significant ($P < 0.001$) genetic differentiation related to the depth. At population level, F_{ST} values between shallow and deep site ranged from 0.051 in Meloria to 0.178 in Piombino. In addition, the differentiation existing between shallow and deep sites in Delimara, Lacco Ameno, Meloria, Marettimo and Stareso was always lower than their differentiation with the shallow and deep sites of other populations (Table 4.2). In Piombino, instead, the differentiation found between shallow and deep sites was higher than the differentiation with the shallow and deep sites in Meloria. Also, the shallow stand in Piombino have a lower distance with the shallow and deep sites in Marettimo and Lacco Ameno than the distance existing with the Piombino deep stand (Table 4.2).

Table 4.2. Pair-wise F_{ST} estimates between shallow and deep sites in the 6 *Posidonia oceanica* populations using neutral markers. DE = Delimara, LA = Lacco Ameno, MA = Marettimo, ME = Meloria, PI = Piombino, ST = Stareso, Sha = shallow site; dee = deep site.

Neutral markers	DE sha	DE dee	LA sha	LA dee	MA sha	MA dee	ME sha	ME dee	PI sha	PI dee	ST sha	ST dee
DE sha	0.000	***	***	***	***	***	***	***	***	***	***	***
DE dee	0.083	0.000	***	***	***	***	***	***	***	***	***	***
LA sha	0.141	0.158	0.000	***	***	***	***	***	***	***	***	***
LA dee	0.178	0.190	0.126	0.000	***	***	***	***	***	***	***	***
MA sha	0.153	0.125	0.145	0.119	0.000	***	***	***	***	***	***	***
MA dee	0.207	0.183	0.179	0.125	0.100	0.000	***	***	***	***	***	***
ME sha	0.233	0.200	0.201	0.203	0.153	0.181	0.000	**	***	***	***	***
ME dee	0.253	0.230	0.224	0.228	0.164	0.192	0.051	0.000	***	***	***	***
PI sha	0.198	0.217	0.172	0.155	0.152	0.175	0.110	0.115	0.000	***	***	***
PI dee	0.349	0.293	0.292	0.320	0.268	0.265	0.160	0.159	0.178	0.000	***	***
ST sha	0.321	0.312	0.304	0.272	0.242	0.220	0.236	0.235	0.218	0.299	0.000	***
ST dee	0.377	0.320	0.350	0.347	0.307	0.302	0.238	0.259	0.280	0.229	0.169	0.000

P-values obtained after 68000 permutations. Indicative adjusted nominal level (5%) for multiple comparisons = 0.000758; ***= P<0.0001; ** = P<0.001.

F_{ST} values calculated with the neutral and selected loci identified along the latitudinal gradient are shown in Table 4.3. The (sub)populations shallow and deep were analyzed separately.

Neutral F_{ST} estimates (Table 4.3) were higher between the Southern Delimara and Lacco (shallow and deep) populations versus all the Northern populations as well for the Northern Stareso deep population versus all the Southern populations. For the remaining populations not always the differentiation versus the populations belonging to the different geographic area was higher than that existing between populations from the same geographic area.

In comparison to the neutral markers, the F_{ST} values calculated with selected markers (Table 4.3) clearly resulted in a higher differentiation between all the southern and northern deep populations than between populations belonging to the same geographic area.

With selected markers, also the Southern shallow populations showed higher differentiation versus the Northern shallow populations, than between populations belonging to the same geographic area, except Delimara shallow population versus Piombino shallow populations.

Table 4.3. Pair-wise F_{ST} between the 12 *Posidonia oceanica* (sub)populations estimated with neutral and selected loci identified along the latitudinal gradient.

Neutral loci along latitudinal gradient								
				South			North	
Shallow populations			DE sha	LA sha	MA sha	ME sha	PI sha	ST sha
	South	DE sha	0.000	***	***	***	***	***
		LA sha	0.137	0.000	***	***	***	***
		MA sha	0.154	0.147	0.000	***	***	***
	North	ME sha	0.239	0.200	0.148	0.000	***	***
		PI sha	0.208	0.163	0.143	0.116	0.000	***
		ST sha	0.303	0.277	0.202	0.248	0.217	0.000
Deep populations			MAdee	LA dee	MA dee	ME dee	PI dee	ST dee
	South	MAdee	0.000	***	***	***	***	***
		LA dee	0.193	0.000	***	***	***	***
		MA dee	0.179	0.118	0.000	***	***	***
	North	ME dee	0.231	0.213	0.165	0.000	***	***
		PI dee	0.275	0.293	0.225	0.167	0.000	***
		ST dee	0.298	0.312	0.248	0.274	0.243	0.000

P-values obtained after 15000 permutations. Indicative adjusted nominal level (5%) for multiple comparisons = 0.003333; ***= P<0.0001.

Selected loci along latitudinal gradient

				South			North	
Shallow populations			DE sha	LA sha	MA sha	ME sha	PI sha	ST sha
	South	DE sha	0.000	***	***	***	**	***
		LA sha	0.174	0.000	***	***	***	***
		MA sha	0.140	0.120	0.000	***	***	***
	North	ME sha	0.179	0.213	0.197	0.000	NS	*
		PI sha	0.093	0.253	0.228	0.056	0.000	***
		ST sha	0.445	0.492	0.494	0.127	0.229	0.000
Deep populations			MAdee	LA dee	MA dee	ME dee	PI dee	ST dee
	South	MAdee	0.000	***	***	***	***	***
		LA dee	0.166	0.000	***	***	***	***
		MA dee	0.215	0.180	0.000	***	***	***
	North	ME dee	0.222	0.329	0.357	0.000	**	**
		PI dee	0.408	0.482	0.491	0.097	0.000	***
		ST dee	0.462	0.544	0.574	0.124	0.099	0.000

P-values obtained after 15000 permutations. Indicative adjusted nominal level (5%) for multiple comparisons = 0.003333; ***= P<0.0001. NS = not significant.

The analysis of molecular variance (AMOVA) with neutral and selected loci is presented in Table 4.4. The analysis with neutral markers identified along the depth gradient was performed considering two groups, the former including all the shallow populations and the latter including all the deep populations. The AMOVA results indicated no significant variation existing between shallow and deep groups of populations. 22 % of total variation was distributed among the populations within the groups and 79% within populations. The AMOVA related to the latitudinal gradient was conducted with selected and neutral loci by defining three different groupings (Table 4.4). Results with neutral markers showed in all the groupings a significant variation among southern and northern populations. With selected loci, variation between the Southern and Northern populations was significantly increased (Table 4.4).

Table 4.4. Analysis of molecular variance (AMOVA) results with neutral and selected microsatellites. For each hierarchical level of analysis (among groups, among populations within groups and within populations) the percentage of variation is reported. Significant differences ($P \leq 0.05$) based on permutation tests (1000 permutations) are marked with asterisks (***), NS = not significant.

	Among groups	Among populations within groups	Within populations
Neutral loci along the depth gradient			
Group I = shallow populations Group II = deep populations	-1.343^{NS}	22.071***	79.272***
Neutral loci along the latitude gradient			
Group I = southern shallow populations Group II = northern shallow populations	4.830***	15.492***	79.678***
Group I = southern deep populations Group II = northern deep populations	7.324***	17.261***	75.414***
Group I = southern shallow and deep population Group II = northern shallow and deep population	6.877***	15.689***	77.434***
Selected loci along the latitude gradient			
Group I = southern shallow populations Group II = northern shallow populations	19.439***	11.757***	68.804***
Group I = southern deep population Group II = northern deep population	34.749***	9.907***	55.344***
Group I = southern shallow and deep population Group II = northern shallow and deep population	28.520***	10.466***	61.014***

The spatial distribution of the (sub)populations (Figure 4.2) resulting from neutral markers along the depth gradient allowed to visualize the lack of a clear separation between the shallow and deep populations, as resulting from AMOVA, but more interestingly the coordinate 1 allowed to visualize a separation between southern and northern populations. This pattern was also better observed when the distribution of the (sub)populations was analyzed with neutral markers along the latitudinal gradient and performed separately for the shallow and deep populations for a best interpretation (Figure 4.2).

Lastly, the PCoA performed with outlier loci identified along the latitudinal gradient (Figure 4.2) showed the clearest separation between the southern and northern groups of shallow and deep populations.

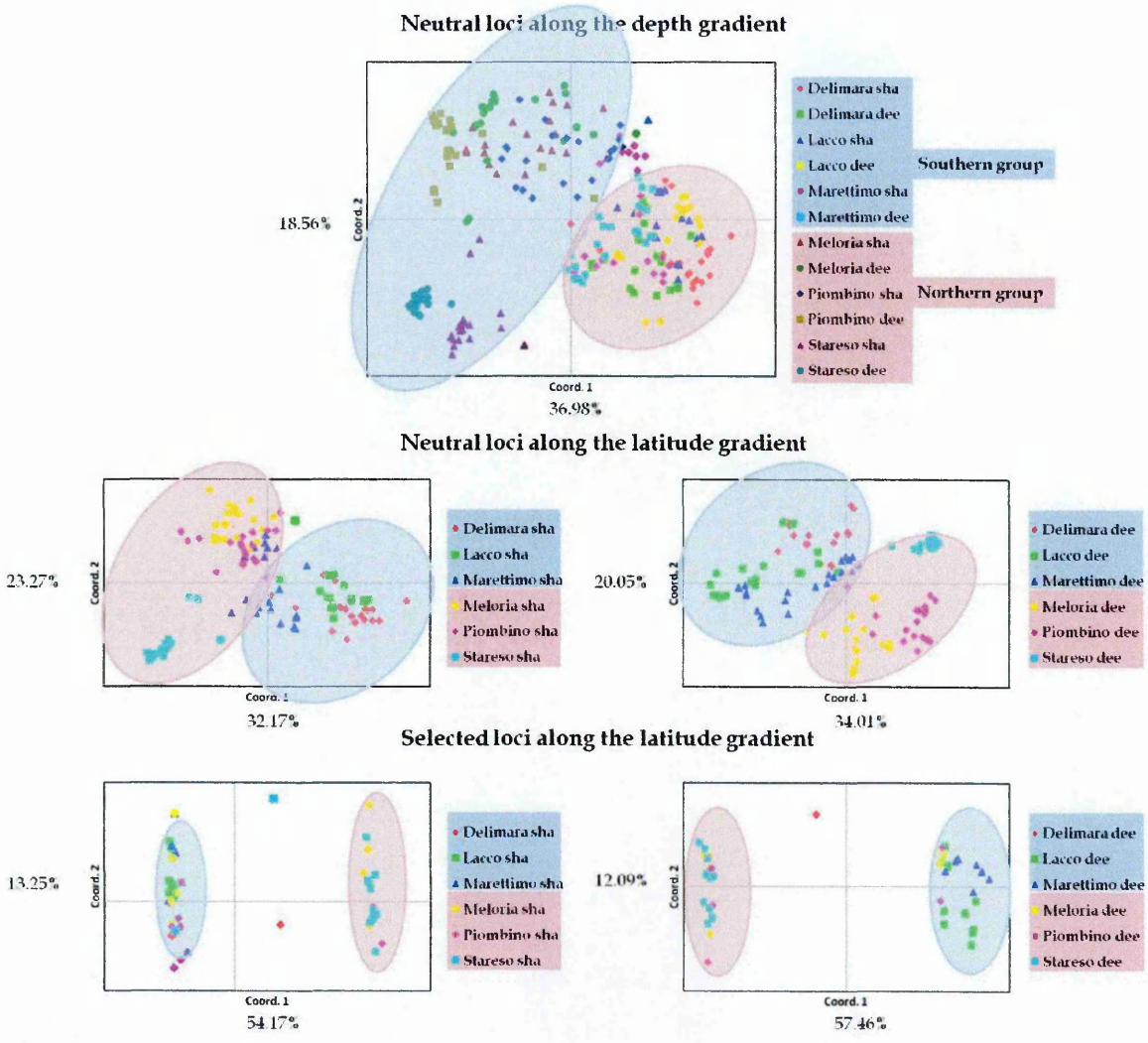


Figure 4.2. Principal coordinate plots calculated with neutral and outlier loci identified along depth and latitudinal gradient. The percentage of variation explained by the coordinates is indicated near the axes. Sha = shallow (sub)population; Dee = deep (sub)population.

4.4.3-*Candidate genes linked to outlier loci*

Candidate genes linked to the outlier loci and putatively involved in adaptive response along the latitudinal gradient were identified: Pooc-153 is involved in the photosynthetic pathway whereas Pooc-PC045G11 is involved in the process of translation (Table 4.5). The outlier locus Poc-*trn* was linked to the *trn*L (UAA) chloroplast intron.

Table 4.5. Putative function of candidate genes associated with light and temperature. In parenthesis are reported KEGG terms.

Microsatellite marker (ID)	Microsatellite position	Gene	KEGG pathway
Poc-trn	*trn*L (UAA) chloroplast intron	–	–
Pooc-153	EST	PsbP	Photosynthesis (K02717)
Pooc-PC045G11	EST	eIF-5A	Translation (K03263)

4.5-Discussion

4-5.1-Signature of selection along a bathymetric and a latitudinal gradient

The genome scan conducted along the latitudinal and depth gradient did not identify the same outlier loci in both analyses. In the genome scan along the bathymetric gradient, no signature of selection was identified, probably as consequence of the limited number of markers used (29 microsatellites). The genome scan performed along the latitude gradient identified 3 outlier loci (10.34 %).

The inconsistency between the two outlier analyses, that were used as proxy to identify signature of selection to temperature and light changes, suggests that the signal detected along the latitudinal gradient may be driven by other environmental factors than temperature or light alone.

The absence of a signal of genetic adaptation along the bathymetric gradient can indicate that either the signature of selection is too weak to be detected with the outliers approach or that a different form of adaptation is taking place. The significant neutral genetic differentiation among populations at different depths suggests limited gene flow between depths that could lead to adaptive genetic differentiation, also supported by differences in the ripening and reproduction periods (Buia and Mazzella, 1991) as well as differences in gene expression (Dattolo et al., 2013; Ruocco et al., 2012) between depths. For example, in Lacco Ameno meadow (Ischia, Gulf of Naples) a difference in *P. oceanica* phenology related to the depth was observed with deep plants showed a delay of about two months in the life cycle in comparison to shallow plants (Buia and Mazzella, 1991).

At the same meadow, gene expression results from a SSH library built with samples from two depths (-5 m and -25 m) showed that transcripts for protein turnover and respiration were more abundant in shallow sites, while transcripts for stress defense, transcription and photosynthesis as well as cell components were predominant in the deep stand (Procaccini et al., 2010). Results from Real Time (qPCR) experiments performed on samples at the two depths showed a down-regulation in the deep plants of transcript of the Photosystem I and II as well as of proteins involved in photo-protection mechanism (Ruocco et al., 2012). Differences in phenology and gene expression suggest that phenotypic plasticity may be involved in the response to light and temperature changes. Indeed, it has been reported that phenotypic plasticity can be a crucial determinant of plant responses, both short- and long-term (Nicotra et al., 2010). In particular, seagrasses display an extraordinarily high degree of phenotypic plasticity that allows them to rapidly respond to changing environmental conditions (Arellano-Méndez et al., 2011). The temperature effects on gene expression have been analyzed in populations of the seagrass *Zostera marina* distributed along a thermal cline (southern and northern populations) as function of their long-term adaptation to heat stress (Franssen et al., 2011). Two salient findings emerged from this work. First, during the heat wave both the northern and southern population showed a similar gene expression characterized by an up-regulation of genes associated with the response to heat (mainly heat shock proteins). Second, shortly thereafter the heat stress (stress recovery) the two populations showed a strong divergence in gene expression. In fact, transcriptomics patterns in southern genotypes returned to control values immediately but genotypes from the northern site failed

to recover and revealed the induction of genes involved in protein degradation, probably reflecting the need to remove from the cell the proteins irreversibly damaged. In conclusion, because the responses to heat-stress were surprisingly similar among genotypes from two locations with widely diverging thermal conditions, the transcriptomic patterns during recovery (transcriptomic resilience) may be a better predictor as to how populations across latitudinal clines are adapted to thermal stress.

The segrasses plasticity to light is a hierarchical process, involving response occurring in leaves, individual shoots and meadow structure (Dalla Via et al., 1998; Dennison, 1987; Longstaff and Dennison, 1999; Olesen et al., 2002). Morphological plasticity at the shoot-scale is a fundamental process that can maximize exposure of the photosynthetic apparatus to light, whilst minimizing respiratory demands. Such morphological adjustments enable the meadow to persist at lower light levels, up to a threshold (Collier et al., 2007). Among the morphological features of seagrasses known to adjust to light availability are canopy height (Bulthius, 1983; West, 1990), leaf width (Lee and Dunton, 1997) and leaf density (Ruiz and Romero, 2001). In particular, under low light conditions in *P. oceanica* it has been observed an increase in the leaf width (Dalla Via et al., 1998). In addition, physiological plasticity has the capacity to prolong seagrass survival under reduced light conditions. It has been shown that at low light seagrass performance is enhanced through a range of acclimative responses (Backman and Barilotti, 1976; Dennison and Alberte, 1986; West, 1990, Olesen and Sand-Jensen, 1993; Abal et al., 1994; Philippart, 1995) leading to altered resource allocation patterns that tend to maximize the efficiency of light harvesting and its conversion

into chemical energy while reducing respiratory costs (Björkman, 1981). The response to light has been well studied in *Halophila stipulacea,* the dominant seagrass in the Gulf of Aqaba (northern Red Sea) where it grows from the intertidal to depths exceeding 50 m (Sharon et al., 2009). This species is a highly plastic seagrass that can acclimate to various light environments quickly. This characteristics was shown for *H. stipulacea* plants grown under high-light conditions, where chloroplast movements were suggested as a means to both capture photons under diurnal low-light periods (when chloroplasts are dispersed) and apparently protect the leaves from photodamage during midday (by clumping the chloroplasts, Sharon and Beer, 2008). The acclimation of this plant to various irradiance along the bathymetric gradient occurred by means of fast changes in the photosynthetic response and light absorption characteristics, which may partly explain its abundance across a wide range of irradiance along the depth gradient that it occupies (Sharon et al., 2009). In addition to the differences in photosynthetic responses at different irradiances, *H. stipulacea* has been reported to feature various leaf morphologies with depth (e.g. longer and wider leaves in deeper water, Lipkin 1979). However, in this case it was not verified whether those changes could be due to acclimation processes or were outcome of longer-term adaptation (Sharon et al., 2009). The study of depth-acclimation in *P. oceanica* and *Cymodocea nodosa* (Olesen et al. 2002) showed that seagrasses can acclimate and respond differently to changing light conditions. Only C. *nodosa,* being a potentially faster growing species compared to *P. oceanica,* adjusted its physiology by reducing photosynthetic light requirements with depth, while both species acclimated mainly at the meadow structure level by reducing

shoot density with the depth. This adjustment should reduce respiratory load and self-shading, enabling greater light penetration through the canopy and is therefore considered a meadow-scale adjustment. When shoot density declines, considerably more light penetrates through the canopy so that the lower sections of leaves growing in deeper water may receive comparable light levels to plants at shallower depths (Dalla Via et al., 1998). The lower morphological and physiological plasticity in response to shading of *P. oceanica,* in comparison to *C. nodosa,* could be explained by the evidence that the two species differ in module size and growth dynamics. *P. oceanica,* in accordance with their slow growth rates of modules and high shoot longevity (Duarte, 1991) is expected to respond more slowly to environmental perturbation, and thus less plastic, compared to *C. nodosa.* The response to light in *Posidonia oceanica* has also been investigated with proteomic approaches (Mazzuca et al., 2009; Serra and Mazzuca, 2011). *Posidonia* meadows acclimated to different light conditions revealed that physiological pathways involved in the acclimation of seagrasses to low light were photosynthetic carbon assimilation, as well as the ubiquitin-dependent proteolysis via proteasomes. In low-light acclimated leaves it was observed a down-regulation of RuBisCo, which is the key regulatory enzyme in the "light-independent reactions" of photosynthesis (Touchette and Burkholder, 2000) and an up-regulation of components of the ubiquitin/mediated proteolysis pathway. Enzymes involved in carbohydrate cleavage were instead up-regulated in low-light conditions, suggesting an increase in the activities to accommodate the demand for metabolic energy during lowlight conditions in leaf tissues. Interestingly, reduced leaf elongation was also observed in aquarium plants

exposed to shading in this study (Mazzuca et al., 2009). In *P. oceanica* meadows, corresponding evidence between RuBisCo down-regulation, and decreased leaf length and shoot density were reported (Acunto et al., 2006). Mazzuca et al. (2009) showed that reduction in leaf growth may be related to the down-regulation of RuBisCo, both in plants acclimated to chronic low-light and in plants exposed to a short periods of shading. In low-light acclimatized chloroplasts, the chla/chlb ratio increases, which is related to the PSII enrichment towards PSI complex (Larkum et al., 2006; Mazzuca et al., 2009). Electron microscopy studies also revealed substantial changes in the stroma lamellae/grana ratios in chloroplasts receiving lowlight, possibly as a mechanism for re-establishing optimal PSI/PSII ratios (Mazzuca et al., 2009).

Adult leaves of *P. oceanica* after two months of artificial shading in aquaria revealed a down-regulation of two chloroplastic carbonic anydrase (CA) in respect to high-light conditions (Serra and Mazzuca, 2011). Seagrasses are, in effect, shade-adapted and they modulate their efficiency in light harvesting by the ability to vary their photosynthetic apparatus to optimize the use of available light (Larkum et al., 2006,). By expressing low levels of choloroplastic CAs, they regulate the uptake of inorganic carbon, thus reaching a new state of carbon assimilation. The authors assumed that the depression of rubisco levels and CA down regulation are related to the maintenance of oprimal PSII/PSI ratios, allowing *P. oceanica* leaves to maximize daily carbon gains under low light conditions (Serra and Mazzuca, 2011).

Recently, physiological acclimation in *Posidonia oceanica* plants along a bathymetric gradient has been analyzed by combining transcriptomic and proteomic analyses

(Dattolo et al., 2013, *in press*). Transcriptional and proteomic profiles showed high differentiation on Chlorophyll a-b binding (Cab) proteins between the two depths. The relative quantity of transcripts and proteins recognized in this study also suggests an increase in PSII and PSI transcripts in deep plants in respect to the shallow ones (especially as regards as PSI). In relation to respiration, an overall increase of related transcripts and proteins was recorded in shallow plants, probably related to the higher temperature present compared with the deeper portion of the meadow plants. Glycolysis and electron transport chain steps were strongly enhanced in high light, while the TCA cycle (tricarboxylic acid cycle) was higher in low light. Several transcripts encoding for proteins associated with stress response and plant defense were detected in low-light. Amongst these, metallothionein-like protein, catalase and oxygen-evolving enhancer proteins and also the cytochromes P450 family. For what concern the protein turnover, E3 ubiquitin-protein ligase, a U-box and RING-box protein and the SCF-E3, F-Box protein appeared to be more expressed in high-light condition in comparison with low-light (Dattolo et al., 2013, *in press*).

Although past research on seagrasses suggest that plasticity may be underlying adaptive responses to light and temperature regimes along a bathymetric gradient, in the present study the inability to identify a clear signature of genetic adaptation to these stressors may be also explained by the limited resolution of our analysis due to the small number of markers used in the genome scan analysis. Genome-wide analysis involving a larger number of markers will help solve this issues.

4.5.2-Candidate gene underlying adaptive response to light and temperature adaptation

The percentage of outlier loci detected in the genome scan along the latitudinal gradient was in line with the information reported in literature. From the genome scan along the latitude gradient, one locus derived from a genomic library (8%) while two loci (12.5%) were linked to EST sequences. In literature, 12–21% of gene-linked markers were detected as under selection in genomic scans in oak species and Atlantic salmon (Scotti-Saintagne et al. 2004; Vasemagi et al. 2005), while 0–9% of anonymous loci showed signature of selection in whitefish, oak, salmon and snails (Wilding et al. 2001; Campbell and Bernatchez 2004; Scotti-Saintagne et al. 2004; Vasemagi et al. 2005). It is already expected that for EST-linked markers more signals of diverging selection can be detected than for genomic markers, because for the latter marker type, chances are higher that markers can reside in intronic and intergenic regions physically distant from a functionally important gene, thus missing the hitch-hiking signal.

Among the candidate genes putatively underlying adaptive response to light and temperature, one gene was involved in the photosynthetic pathway, and one in the process of translation.

The first candidate gene identified, PsbP, is associated with the oxygen-evolving enhancer protein 2, a component of the photosystem II (PSII). This protein is part of the oxygen evolving complex (OEC or water splitting complex) involved in the photo-oxidation of water during the light reactions of photosynthesis. PsbP is required for PSII to be fully operational *in vivo* because it increases the affinity of the water oxidation site for chloride ions and provides the conditions required for

high affinity binding of calcium ions (Kochhar et al., 1996). Recent RNAi experiments in *Arabidopsis thaliana* have shown that this lumen protein is both essential and quantitatively related to PSII efficiency and stability (Ishihara et al., 2005; Ifuku et al., 2005). The identification of a component of photosystem II, and in particular a member of the OEC, makes much sense as gene underlying response to light changes, as the exposure of photosynthetic organisms to strong light results in inhibition of the activity of PSII (Powles, 1984; Prásil et al., 1992; Aro et al. 1993; Andersson and Aro, 2001). This phenomenon, referred to as photoinhibition, is overcome through the rapid and efficient repair of PSII (Prásil et al., 1992; Aro et al. 1993; Andersson and Aro, 2001),

In addition, it has been shown that the activity of PSII declines under unfavorable or stressful environmental conditions also presented by high salt concentration, (Sharma and Hall, 1991; Hertwig et al., 1992), low and moderate temperatures (Öquist and Huner 1991; Öquist et al., 1993; Berry, J. and Björkman, 1980) and CO_2 limitation (Sicher, 1984; Miller, 1989). Differently from the light that acts directly on the OEC, these latter factors act primarily by inhibiting the repair of PSII by suppressing the synthesis of proteins *de novo* required for the repair of PSII (Murata et al., 2007). Ishihara et al. (2005), showed that in *N. tabacum* PSII activity was linearly correlated with the total amounts of PsbP expressed and suggested that PsbP over-expression could improve electron flow through PSII (Ishihara et al., 2005; Ifuku et al., 2005). In particular, it has been shown that under biotic and abiotic stresses in various plant species the amount of PsbP changes: the expression of PsbP was regulated by environmental stimuli such as salt, drought and virus infection in rice (Abbasi and Komatsu, 2004) and *Nicotiana benthamiana*

(Pérez-Bueno et al., 2004). In *N. tabacum*, differential regulation of the expression of PsbP has been reported during the progression of viral infection (Takahashi et al., 1991; Takahashi and Ehara, 1992). The differential accumulation of PsbP has also been shown also as adaptive response to high-salt concentration in cultured tobacco cells (Murota et al., 1994; Sato et al., 1995). Interestingly, EST encoding for PsbP are over-represented in wheat plants exposed to cold stress (Houde et al., 2006)

It is possible that the action of selection on this component of the photosystem II in *P. oceanica* allows the plant to respond to changes in light and temperature that could directly or indirectly damage the photosystem II. By maintaining a flexible modulation of PSII activity the plants could improve the electron flow through PSII and prevent its decline.

The second candidate gene identified in this study, eIF-5A-1, is associated to an eukaryotic translation initiation factor 5A-1. Plant eIF5A proteins are highly conserved and are involved in multiple biological processes, including gene regulation, translation elongation, mRNA turnover and decay, cell proliferation, leaf and root growth, seed yield, leaf, flower and fruit senescence and programmed cell death (Wang et al., 2003; Ma et al., 2010a; Xu et al., 2011). Ma et al. (2010a) showed that eIF5A plays roles in supporting plant growth and in regulating responses to sub-lethal osmotic and nutrient stress (Ma et al., 2010b). Valentini et al. (2002) showed that eIF5A is involved in the WSC/PKC1 signaling pathway that controls cell wall integrity or related processes, and plays a role in cell wall formation (Valentini et al., 2002). Hopkins et al. reported that eIF5A plays a vital role in signal transduction pathways involved in pathogen-induced cell

death and in the development of plant disease symptoms. Interestingly, plant eIF5A genes are also involved in abiotic stress responses (Hopkins et al., 2008). In *Rosa chinensis* the gene that encodes an eIF5A (*Rce*IF5A) is up-regulated under high temperature, and oxidative and osmotic stress conditions (Xu et al., 2011). Over-expression of *Rce*IF5A in transgenic *Arabidopsis* resulted in plants exhibiting strong resistance to high temperature stress. In addition, superoxide dismutase (SOD) activity was more enhanced in *Rce*IF5A overexpressing plants after oxidative stress, which demonstrates that *Rce*IF5A over-expression enhanced oxidative stress resistance by regulating antioxidant enzyme activities, and suggests that the accumulation of *Rce*IF5A in vivo might increase SOD activity by participating in recruitment or translation of mRNAs from genes responsible for the SOD synthesis pathway. RT-PCR analysis revealed that there is also a connection between RceIF5A and the expression of two other antioxidant enzyme genes (catalase and gluthatione peroxidase) in transgenic *Arabidopsis*. This strongly supported that *Rce*IF5A plays such an important role in heat/oxidative stress responding pathway and participates in the antioxidant enzyme synthesis pathway. The proline content of plants under osmotic stress was also calculated and the result indicated that *Rce*IF5A over-expression enhanced osmotic stress resistance by affecting the proline synthesis pathway (Xu et al., 2011). Chou et al. 2004 reported that salt and heavy metal stress induce the expression of rice eIF5A genes, *Ose*IF5A-1 and *Ose*IF5A-2, suggesting that these genes are involved in stress tolerance (Chou et al., 2004). Wang et al. (2012) suggested that eIF5A may play an important role in plant adaptation to changing environmental conditions. In *Tamarix androssowii*, *Tae*IF5A1 may mediate abiotic stress tolerance by increasing

protein synthesis, enhancing ROS scavenging by improving SOD and peroxidase (POD) activities, and preventing chlorophyll loss and membrane damage. These results suggested that eIF5A is involved in eliciting a stress response mechanism that may play a common role in plant tolerance to salt, heat, oxidative, osmotic and heavy metal stresses (Wang et al., 2012). The identification of this gene as outlier locus in *P. oceanica* suggests that its role in the response to temperature and light changes may be critical to activate the translation of protective molecules against the oxidative stress and ensure the survival under adverse environmental conditions.

The microsatellite Poc-*trn* identified as outlier locus was linked to a non coding region, the intron internal to the tRNA leucine gene (trnL). This intron is peculiar in being the only Group I intron in the chloroplast DNA (Taberlet et al., 2007), a class of ribozymes that catalyze their own excision (self-splicing) from the flanking exons ribozyme (Nielsen and Johansen, 2009). Evidence is lacking on *in vivo* function and biological role despite the fact that this intron is highly abundant in nature. The identification of this intron as locus under selection could suggest a putative biological role in relation to light and temperature changes, possibly through its ribozyme functionality that may regulate tRNA activity as well as protein-coding gene expression.

4.6-Conclusions

This study was a first step towards the identification of candidate genes underlying adaptive responses to light and temperature regimes in the seagrass *Posidonia oceanica*. Along the depth gradient, no loci were detected as outliers probably due to technical limitation related to the low number of loci used. Alternatively, phenotypic plasticity appears to be playing a stronger role than genetic adaptation in response to these environmental changes. Along the depth gradient, 3 loci were identified to be under selection. The fact that these loci were identified only along a latitudinal gradient may suggest that the detected signature of selection underlies adaptation to other environmental challenges than temperature and/or light alone. Further studies will help elucidating the relative role of genetic adaptation and phenotypic plasticity in this keystone species.

4.7-References

Abal E. G., Loneragan N., Bowen P., Perry C. J., Udy J. W., Dennison W. C. (1994). Physiological and morphological responses of the seagrass *Zostera carpricorni* Aschers. to light intensity. *J Exp Mar Biol Ecol* 178: 113–119.

Abbasi, F. M. and Komatsu, S. (2004). A proteomic approach to analyze saltresponsive proteins in rice leaf sheath. *Proteomics* 4: 2072–2081.

Acunto, S., Rende, F., Cozza, R., Innocenti, A.M., (2006). State of health of *Posidonia oceanica* meadows along the northern Tyrrhenian coast of Calabria, Italy. *Biol. Mar. Medit.* 13: 8–13.

Alberto, F., Correia, L., Arnaud-Haond, S., Billot, C., Duarte, C. M., & Serrao, E. (2003). New microsatellite markers for the endemic Mediterranean seagrass *Posidonia oceanica. Molecular Ecology Notes*, 3(2): 253–255.

Andersson, B., Aro, E.M. (2001). Photodamage and D1 protein turnover in photosystem II, in: E.M. Aro, B. Anderson (Eds.), Regulation of Photosynthesis, Kluwer Academic, Dordrecht, The Netherlands, pp. 377–393.

Antao T., Lopes, A., Lopes R. J., Beja-Pereira, A., Luikart, G. (2008). LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics*, 9: 323.

Arellano-Méndez, L. U., Herrera-Silveira, J. A., Montero- Muñoz, J. L., Liceaga-Correa, M. A. (2011). Morphometric trait variation in *Thalassia testudinum* (Banks ex Konig) associated to environmental heterogeneity in a subtropical ecosystem. *J Ecosys Ecograph*, 1: 1-13.

Aro, E. M., Virgin, I., Andersson, B. (1993). Photoinhibition of photosystem II. Inactivation, protein damage and turnover. *Biochim. Biophys. Acta* 1143: 113–134.

Backman, T. W., Barilotti, D. C. (1976). Irradiance reduction: effects on standing crops of the eelgrass Zostera marina in a coastal lagoon. *Mar Biol* 34: 33–40.

Beaumont, M. A., and R. A. Nichols. (1996) Evaluating loci for use in the genetic analysis of population structure. Proc. R. Soc. Lond. *B Biol. Sci.* 263: 1619–1626.

Beer, S. and E. Koch. (1996). Photosynthesis of seagrasses vs. marine macroalgae in globally changing CO_2 environments. *Marine Ecology Progress Series* 141: 199-204.

Berry, J. and O. Björkman. (1980). Photosynthetic response and adaptation to temperature in higher plants, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 31: 491–543.

Björk, M., Short, F.T., McLeod, E., Beer, S. (2008). Managing Seagrasses for Resilience to Climate Change. IUCN, Gland, Switzerland.

Björkman, O. (1981) Responses to different quantum flux densities. In: Lange OL, Nobel PS, Osmond CB, Ziegler H (eds) Encyclopaedia of plant physiology, Vol 12A. Springer-Verlag, Berlin, p 57–107

Bonin, A., Taberlet, P., Miaud, C., Pompanon, F. (2006). Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Mol Biol Evol* 23: 773–783.

Buia, M. C. and Mazzella L. (1991). Reproductivity phenology of the Mediteranean Seagrasses *Posidonia oceanica* (L) Delile *Cymodocea nodosa* (Ucria) *Aschers Zostera noltii Hornem. Aquat. Bot.*, 40(1): 343-362.

Bulthuis, D. A.(1983). Effects of temperature on the photosynthesis- irradiance curve of the Australian seagrass, *Heterozostera tasmanica. Mar. Biol. Lett.* 4: 47–57.

Campbell, D., and L. Bernatchez. (2004). Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Mol. Biol. Evol.* 21: 945–956.

Chou, W. C., Huang, Y. W., Tsay, W. S., Chiang, T. Y., Huang, D. D., Huang, H. J. (2004). Expression of genes encoding the rice translation initiation factor, eIF5A, is involved in developmental and environmental responses. *Physiol Plant*, 121(1): 50–57.

Collier, C. J., Lavery, P. S., Masini, R.J., Ralph, P. J. (2007). Morphological, growth and meadow characteristics of the seagrass Posidonia sinuosa along a depth-related gradient of light availability. *Mar. Ecol. Prog. Ser.*, 337, 103–115.

Dalziel, A. C, Rogers, S. M, Schulte, P. M. (2009). Linking genotypes to phenotypes and fitness: how mechanistic biology can inform molecular ecology. *Molecular Ecology*, 18: 4997–5017.

Dalla Via, J., Sturmbauer, C., Schonweger, G., Sotz, E., Mathekowitsch, S., Stifter, M., Rieger, R. (1998). Light gradients and meadow structure in *Posidonia oceanica*: ecomorphological and functional correlates. *Mar. Ecol. Prog. Ser.*, 163: 267–278.

Dattolo, E., Gu, J., Mazzuca, S., Serra, I.A., Spadafora, A., and Procaccini, G. (2013) *in press.* Transcriptional and proteomics profiles of Posidonia oceanica (seagrass) along its bathymetrical distribution. *Frontiers in Plant proteomics.*

Dennison, W. C., and R.S., Alberte. (1986). Photoadaptation and growth of *Zostera marina* L. (eelgrass) transplants along a depth gradient. *J Exp Mar Biol Ecol*, 98: 265–282.

Dennison, W. C.(1987). Effects of light on seagrass photosynthesis, growth and depth distribution. *Aquat. Bot.*, 27: 15–26.

Dieringer, D., Schlötterer, C. (2003). Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* 3: 167-169.

Duarte, C. M. (1991). Seagrass depth limits. *Aquatic Botany*, 40(4): 363–377.

Duarte, C. M., Borum, J., Short, F. T., Walker, D. I. (2008). Seagrass ecosystems: their global status and prospects. In: Polunin, N.V.C. (Ed.), Aquatic Ecosystems: Trends and Global Prospects. Cambridge University Press, pp, pp. 281–294.

Excoffier, L., Laval, G., Schneider, S. (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1: 47-50.

Foll, M., and O. E., Gaggiotti. (2008). Estimating selection with different markers and varying demographic scenarios: a Bayesianperspective. *Genetics*, 180: 977-993.

Franssen, S. U., Gu, J., Bergmann, N., Winters, G., Klostermeier, U.C., Rosenstiel, P., Bornberg-Bauer, E., Reusch, T. B. H. (2011) Transcriptomic resilience to global warming in the seagrass *Zostera marina*, a marine foundation species. *Proceedings of the Natural Academy of Sciences of the United States of America*, 108: 19276-19281.

Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3).

Hansen, M. M., Olivieri, I., Waller, D. M., Nielsen, E. E. (2012). Monitoring adaptive genetic responses to environmental change. *Molecular ecology*, 21(6): 1311-29.

Hansen, J., (2007). Climate catastrophe. *New Scientist*, 195: 30-34.

Harley, C. D. G., A. R. Hughes, K. M. Hultgren, B. J. Miner, C. J. B. Sorte, C. S. Thornber, L. F. Rodriguez, L. Tomanek, and Williams, S.L. (2006). The impacts of climate change in coastal marine systems. *Ecology Letters*, 9: 228-241.

Hertwig, B., Streb, P., Feierabend, J. (1992). Light dependence of catalase synthesis and degradation in leaves and the influence of interfering stress conditions, *Plant Physiol.*, 100: 1547-1553.

Hoffmann, A., and T., Willi (2008). Detecting genetic response to environmental change. *Nature Reviews Genetics*, 9: 421-432.

Holderegger, R., Herrmann, D., Poncet, B. et al. (2008). Land ahead: using genome scans to identify molecular markers of adaptive relevance. *Plant Ecology & Diversity*, 1: 273-283.

Hopkins, M. T, Lampi, Y., Wang, T. W., Liu, Z., Thompson, J. E (2008). Eukaryotic translation initiation factor 5A is involved in pathogen-induced cell death and development of disease symptoms in Arabidopsis. *Plant Physiol*, 148(1): 479-489.

Houde, M., Belcaid, M., Ouellet, F., Danyluk, J., Monroy, A.F., Dryanova, A., Gulick, P., Bergeron, A., Laroche, A., Links, M.G. et al. (2006). Wheat EST resources for functional genomics of abiotic stress. *BMC Genomics*, 7: 149.

Ifuku, K., Yamamoto, Y., Ono T. A., Ishihara, S., Sato, F. (2005). PsbP protein, but not PsbQ protein, is essential for the regulation and stabilization of photosystem II in higher plants. *Plant Physiol*, 139: 1175-1184.

Intergovernmental Panel on Climate Change (IPCC). (2007). Climate Change 2007: The Physical Science Basis. Summary for Policy Makers. 10th Session of the Working Group I of the IPCC. Paris, France.

Ishihara, S., Yamamoto, Y., Ifuku, K., Sato, F. (2005). Functional analysis of four members of the PsbP family in photosystem II in Nicotiana tabacum using differential RNA interference. *Plant Cell Physiol*, 46: 1885-1893.

Jeffreys, H. (1961). Theory of Probability . 3rd ed. Oxford Classic Texts in the Physical Sciences, Oxford: Oxford University Press

Jump, A. S., Hunt, J. M., Martínez-Izquierdo, J. A.,Peñuelas, J. (2006). Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. *Molecular ecology*, 15(11): 3469–80.

Kauer, M. O., Dieringer, D., Schlötterer, C. (2003). A microsatellite variability screen for positive selection associated with the "out of Africa" habitat expansion of *Drosophila melanogaster*. *Genetics*, 165(3): 1137–48.

Kawecki, T. J., Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7(12): 1225–1241.

Kochhar, A., Khurana, J. P., Tyagi, A. K. (1996). Nucleotide sequence of the psbP gene encoding precursor of 23-kDa polypeptide of oxygen-evolving complex in *Arabidopsis thaliana* and its expression in the wild-type and a constitutively photomorphogenic mutant. *DNA Res.* 31; 3(5): 277-85.

Larkum, W. D., Orth, R. J, Duarte, C. M., eds. (2006). Seagrasses: Biology, Ecology and Conservation. Dordrecht (The Netherlands): Springer.

Lee, K., Dunton, K.H., (1997). Effects of in situ light reduction on the maintenance, growth, and partitioning of carbon resources in *Thalassia testudinum* Banks ex König. *J. Exp. Mar. Biol. Ecol.* 210: 53–73.

Lipkin, Y. (1979). Quantitative aspects of seagrass communities, particularly of those dominated by *Halophila stipulacea*, in Sinai (Northern Red Sea). *Aquat Bot* , 7: 119–128

Longstaff, B.J., Dennison, W.C. (1999). Seagrass survival during pulsed turbidity events: the effects of light deprivation on the seagrasses *Halodule pinifolia* and *Halophila ovalis*. *Aquat. Bot,*. 65: 101–121.

Luikart, G., England, P. R., Tallmon, D., Jordan, S., Taberlet, P. (2003). The power and promise of population genomics: from genotyping to genome typing. *Nature reviews. Genetics*, 4(12): 981–94.

Ma, Y., Miura, E., Ham, B. K., Cheng, H. W., Lee, Y. J, Lucas, W. J. (2010a). Pumpkin eIF5A isoforms interact with components of the translational machinery in the cucurbit sieve tube system. *The Plant Journal*, 64(3): 536–550.

Ma, F., Liu, Z., Wang, T. W. E.I., Hopkins M. T., Peterson, C. A., Thompson, J. E.(2010b). *Arabidopsis* eIF5A3 influences growth and the response to osmotic and nutrient stress. *Plant Cell Environ*, 33(10): 1682–1696.

Mazzuca, S., Filadoro, D., Vannini, C., Marsoni, M., Cozza, R., Bracale, M., Pangaro, T., and Innocenti A.M. (2009). Seagrass light acclimation: 2-DE protein analysis in *Posidonia* leaves grown in chronic low light conditions. *Journal of Experimental Marine Biology and Ecology*, 374: 113–122.

Maynard, S. J. and J. Haigh. (1974). The hitch-hiking effect of a favourable gene. *Genetical Research,* 23: 23- 35.

Mealor, B. A., and A.L .Hild. (2006). Potential selection in native grass populations by exotic invasion. *Molecular Ecology,* 15: 2291-2300.

Meyer, C. L., Vitalis, R., Saumitou-Laprade, P., Castric, V. (2009). Genomic pattern of adaptive divergence in *Arabidopsis halleri* , a model species for tolerance to heavy metal. *Molecular Ecology,* 18(9): 2050-2062.

Migliaccio, M., De Martino, F., Silvestre, F., Procaccini, G. (2005). Meadow-scale genetic structure in *Posidonia oceanica. Marine Ecology Progress Series,* 304: 55-65.

Miller, A. G. and D. T. Canvin.. (1989). Glycolaldehyde inhibits CO_2 fixation in the cyanobacterium *Synechococcus* UTEX 625 without inhibiting the accumulation of inorganic carbon or the associated quenching of chlorophyll a fluorescence. *Plant Physiol.,* 91: 1044-1049.

Murata, N., Takahashi, S., Nishiyama, Y., Allakhverdiev, S. I .(2007). Photoinhibition of photosystem II under environmental stress. *Biochim Biophys Acta,* 1767: 414-421-

Murota, K., Ohshita, Y., Watanabe, A., Aso, S., Sato, F.,Yamada, Y. (1994). Changes related to salt tolerance in thylakoid membranes of photoautotrophically cultured green tobacco cells. *Plant Cell Physiol.,* 35: 107-113.

Namroud, M. C., Beaulieu J., Juge N., Laroche J., Bousquet, J. (2008). Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. *Mol Ecol,* 17: 3599-3613

Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E. J., Mathesius, U., Poot, P., Purugganan, M. D. , Richards, C. L., Valladares, F., van Kleunen, M. (2010). Plant phenotypic plasticity in a changing climate. *Trends in Plant Science,* 15: 684-692.

Nielsen, H. and Johansen, S. D. (2009). Group I introns - moving in new directions. *RNA Biology,* 6 (4): 1-9.

Oetjen, K., and T. B. H. Reusch. (2007). Genome scans detect consistent divergent selection among subtidal vs. intertidal populations of the marine angiosperm *Zostera marina. Molecular ecology,* 16(24): 5156-67.

Oetjen, K., Ferber, S., Dankert, I., & Reusch, T. B. H. (2009). New evidence for habitat-specific selection in Wadden Sea *Zostera marina* populations revealed by genome scanning using SNP and microsatellite markers. *Marine Biology,* 157(1): 81-89.

Olesen, B. andK. Sand-Jensen (1993). Seasonal acclimatization of eelgrass *Zostera marina* growth to light. *Mar Ecol Prog Ser,* 94: 91-99.

Olesen, B., et al., (2002). Depth acclimation of photosynthesis, morphology and demography of Posidonia oceanica and *Cymodocea nodosa* in the Spanish Mediterranean Sea. *Mar. Ecol. Prog. Ser.* ,. 236: 89-97.

Öquist, G., Huner, N.P.A. (1991). Effects of cold acclimation on the susceptibility of photosynthesis to photoinhibition in Scots pine and in winter and spring cereals: a fluorescence analysis. *Funct. Ecol.*, 5: 91-100.

Öquist, G., Hurry, V.M., Huner, N.P.A. (1993). Low-temperature effects on photosynthesis and correlation with freezing tolerance in spring and winter cultivars of wheat and rye. *Plant Physiol.*, 101: 245-250.

Orth, R. J, Carruthers, T. J. B, Dennison, W. C, Duarte, C. M., Fourqurean, J. W., JR, K. L. H., Hughes, A. R, Olyarnik, S., Williams, S. L., Kendrick, G. A., Kenworthy, W. J., Short, F. T., Waycott, M. (2006). A Global Crisis for Seagrass Ecosystems. BioOne, 56(12): 987-996.

Overpeck, J. T., Otto-Bliesner, B. L., Miller, G. H., Muhs, D. R., Alley, R. B. and Kiehl, J. T. (2006). Paleoclimatic evidence for future ice-sheet instability and rapid sea-level rise. *Science*, 311: 1747-1750.

Peakall, R., and P. E. Smouse.(2006). Genalex 6: genetic analysis in Excel. Population geneticsoftware for teaching and research. *Molecular Ecology Notes*, 6(1): 288-295.

Pérez-Bueno, M. L., Rahoutei, J., Sajnani, C., García-Luque, I. and Barón, M. (2004). Proteomic analysis of the oxygen-evolving complex of photosystem II under biotic stress: studies on *Nicotiana benthamiana* infected with tobamoviruses. *Proteomics*, 4: 418-425.

Philippart, C. J. M. (1995). Effects of shading on growth, biomass and population maintenance of the intertidal seagrass *Zostera noltii* Hornem. in the Dutch Wadden Sea. *J Exp Mar Biol Ecol* 188: 199-213

Powles, S. B. (1984). Photoinhibition of photosynthesis induced by visible light. *Annu. Rev. Plant Physiol.*, 35: 15-44.

Prásil, O., Adir, N., Ohad, I. (1992). Dynamics of photosystem II: mechanism of photoinhibition and recovery process, in: J. Barber (Ed.), *The Photosystems: Structure, Function and Molecular Biology*, vol. 11, Elsevier Science Publishers, Amsterdam, The Netherlands, pp. 295-348.

Procaccini, G., and Waycott, M. (1998). Microsatellite loci identified in the seagrass *Posidonia oceanica* (L.) Delile. *Journal of Heredity*, 89. 562-568.

Procaccini, G., Dattolo, E., Innocenti, AM., Mazzuca, S., Cavallini, A., Natali, L.J., Bayer, P.G. (2010). Differential gene expression profiling under different light conditions in *Posidonia oceanica* (L.) Delile by SSH analysis. *Rapp Comm Int Mer Médit*, 39: 641

Ralph, P. J., Durako, M. J., Enríquez, S., Collier, C. J., Doblin, M. A. (2007). Impact of light limitation on seagrasses. *Journal of Experimental Marine Biology and Ecology*, 350(1-2). 176-193.

Rahmstorf, S. (2007). A Semi-empirical approach to projecting future sea-level rise. *Science*, 315: 368-370.

Reusch, T. B. H. and Wood, T. E. (2007). Molecular ecology of global change. *Mol. Ecol,*. 16. 3973-3992.

Ruocco, M., Brunet, C., Lorenti, M., Lauritano, C., D'Esposito, D., Riccio, M., Procaccini, G. (2012). *Posidonia oceanica* photoadaptation to the depth gradient. *Biol. Mar. Mediterr.*, 19 (1): 63-64.

Ruiz, J.M., and J. Romero. (2001). Effects of in situ experimental shading on the Mediterranean seagrass *Posidonia oceanica. Mar. Ecol. Prog. Ser.*,215: 107-120.

Sato, F., Kanda, Y., Shiga, M., Ohshita, Y., Murota, K. and Yamada, Y. (1995). Characterization of OEC23 polypeptide of NaCl-adapted photoautotrophic cultured cells of tobacco. In Photosynthesis: From Light to Biosphere. Edited by P. Mathis. pp. 601-604. Kluwer Academic Publishers, The Netherlands.

Savolainen, O., Pyhäjärvi, T.,Knürr, T. (2007). Gene Flow and Local Adaptation in Trees. Annual *Review of Ecology, Evolution, and Systematics*, 38(1). 595-619.

Schlötterer, C. (2002a). A microsatellite-based multilocus screen for the identification of local selective sweeps. *Genetics*, 160: 753-763.

Schlötterer, C. (2002b). Towards a molecular characterization of adaptation in local populations. *Curr. Opin. Genet. Dev.*, 12: 683-687.

Schlötterer, C. and D. Dieringer (2005). A novel test statistics for the identification of local selective sweeps based on microsatellite gene diversity. Selective sweeps Eurekah.com and Klüver Academic/Plenum Publishers, Georgetown, TX, USANurminski D, 55-64.

Scotti-Saintagne, C., Mariette, S., Porth, I. et al. (2004). Genome scanning for interspecific differentiation between two closely related oak species *Quercus robur* L. & *Q petraea* (Matt.) Liebl. *Genetics*, 168: 1615-1626.

Serra, I.A. and S.Mazzuca. (2011). *Posidonia oceanica*: from ecological status to genetic and proteomic resources, In: Seagrass: Ecology, Uses and Threats, ed. R.S. Pirog (Nova Science Publishers, New York), 71-116.

Sharma, P.K. and D.O. Hall. (1991). Interaction of salt stress and photoinhibition on photosynthesis in barley and sorghum. *J. Plant Physiol,*. 138: 614-619.

Sharon, Y.,and S. Beer.(2008). Diurnal movements of chloroplasts in *Halophila stipulacea* and their effect on PAM fluorometric measurements of photosynthetic rates. *Aquat Bot*, 88: 273-276

Sharon, Y., Silva, J., Santos, R., Runcie, J. W., Chernihovsky, M.,Beer, S. (2009). Photosynthetic responses of *Halophila stipulacea* to a light gradient. II. Acclimations following transplantation. *Aquat. Bot.*, 7: 153-157

Sicher, R.C. (1984). Glycolaldehyde inhibition of photosynthetic carbon assimilation by isolated chloroplasts and protoplasts, in: C. *Sybesma* (Ed.), Advances in Photosynthetic Research, vol. 3, W. Junk, Hague, pp. 413-416.

Short, F. T. and H.A. Neckles. (1998). The effects of global climate change on seagrasses. *Aquatic Botany*,63: 169-196.

Short, F. T., Coles, R.G., Pergent-Martini, C.(2001). Global seagrass distribution. In: Short, F.T., Coles, R.G. (Eds.), Global Seagrass Research Methods. Elsevier Science B.V, Amsterdam, pp. 5-30.

Storz, J. F. (2005). Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol. Ecol.*, 14: 671-688.

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermat, T., Corthier, G., Brochmann, C., Willerslev, E. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, 35: 8.

Takahashi, H., Ehara, Y.,Hirano, H. (1991). A protein in the oxygen-evolving complex in the chloroplast is associated with sympton expression on tobacco leaves infected with cucumber mosaic virus strain Y. *Plant Mol. Biol.*,16: 689-698.

Takahashi, H. and Y. Ehara. (1992). Changes in the activity and the polypeptide composition of the oxygen-evolving complex in photosystem II of tobacco leaves infected with cucumber mosaic virus strain Y. *Mol. Plant Microbe Interact.*, 5: 269-272.

Tang, J., Hanage, W. P., Fraser, C., Corander, J. (2009). Identifying currents in the gene pool for bacterial populations using an integrative approach. *PLoS Comput Biol*, 5:e1000455.

Touchette, B. W., Burkholder, J. M. (2000). Overview of physiological ecology of carbon metabolism in seagrasses. *J. Expert. Mar. Biol. Ecol.*, 250: 163-205.

Trenberth, K. (2005). Uncertainty in hurricanes and global warming. *Science*, 308: 1753-1754.

Valentini, S. R., Casolari, J. M., Oliveira, C. C., Silver, P. A., McBride, A. E. (2002). Genetic interactions of yeast eukaryotic translation initiation factor 5A (eIF5A) reveal connections to poly (A)-binding protein and protein kinase C signaling. *Genetics*, 160(2): 393.

Vasemägi, A., Nilsson, J., Primmer, C. R. (2005). Expressed sequence Tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic Salmon (*Salmo salar* L.). *Mol Biol Evol*, 22: 1067-1076.

Xu, J., Zhang, B., Jiang, C., Ming, F, (2011). RceIF5A, encoding an eukaryotic translation initiation factor 5A in *Rosa chinensis*, can enhance thermotolerance, oxidative and osmotic stress resistance of *Arabidopsis thaliana*. *Plant Mol Biol*, 75: 167-178.

Wang, T. W, Lu, L., Zhang, C. G., Taylor, C., Thompson, J. E. (2003). Pleiotropic effects of suppressing deoxyhypusine synthase expression in Arabidopsis thaliana. *Plant Mol Biol*, 52(6): 1223-1235.

Wang, T., Chen, G., Zan, Q., Wang, C., Su, Y. (2012). AFLP genome scan to detect genetic structure and candidate loci under selection for local adaptation of the invasive weed Mikania *micrantha*. *PloS one*, 7(7), e41310.

Waycott, M., Duarte, C. M., Carruthers, T. J. B., Orth, R. J., Dennison, W. C., Olyarnik, S., Calladine, A., et al. (2009). Accelerating loss of seagrasses across the globe threatens coastal ecosystems. Proceedings of the National Academy of Sciences of the United States of America, 106(30), 12377-81. doi:10.1073/pnas.0905620106

West, R.J. (1990). Depth-related structural and morphological variations in an Australian *Posidonia* seagrass bed. *Aquat Bot,* 36: 153-166

Wilding, C. S., Butlin, R. K., Grahame, J. (2001). Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *Journal of Evolutionary Biology,* 14: 611-619.

4.8-Appendix

Summary of outlier analysis performed along the bathymetric and latitudinal gradient

Following are reported the visualization of the polymorphic loci involved in the pair-wise comparisons (Figure 4.A1) and examples of the outputs from the tree outlier tests adopted (Plot 4.A1, Plot 4.A2,Plot 4.A3, Table 4.A1, Table 4.A2). The table 4.A3 and 4.A4 summarize the overall results obtained from the outlier tests that have been used for the calculation of outliers frequency distribution reported in Figure 4.A2.

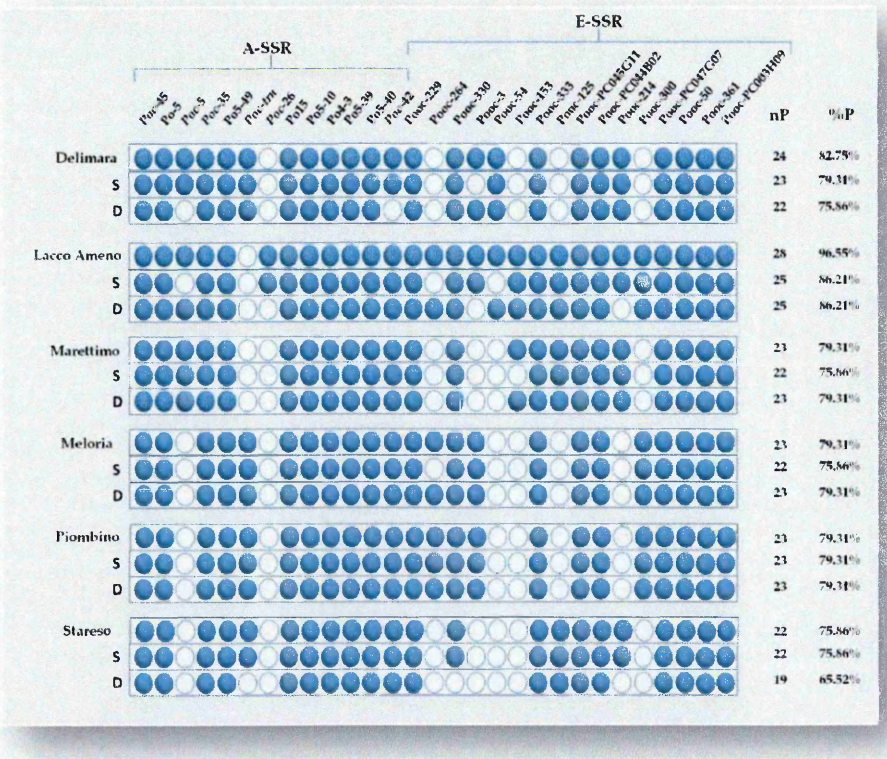
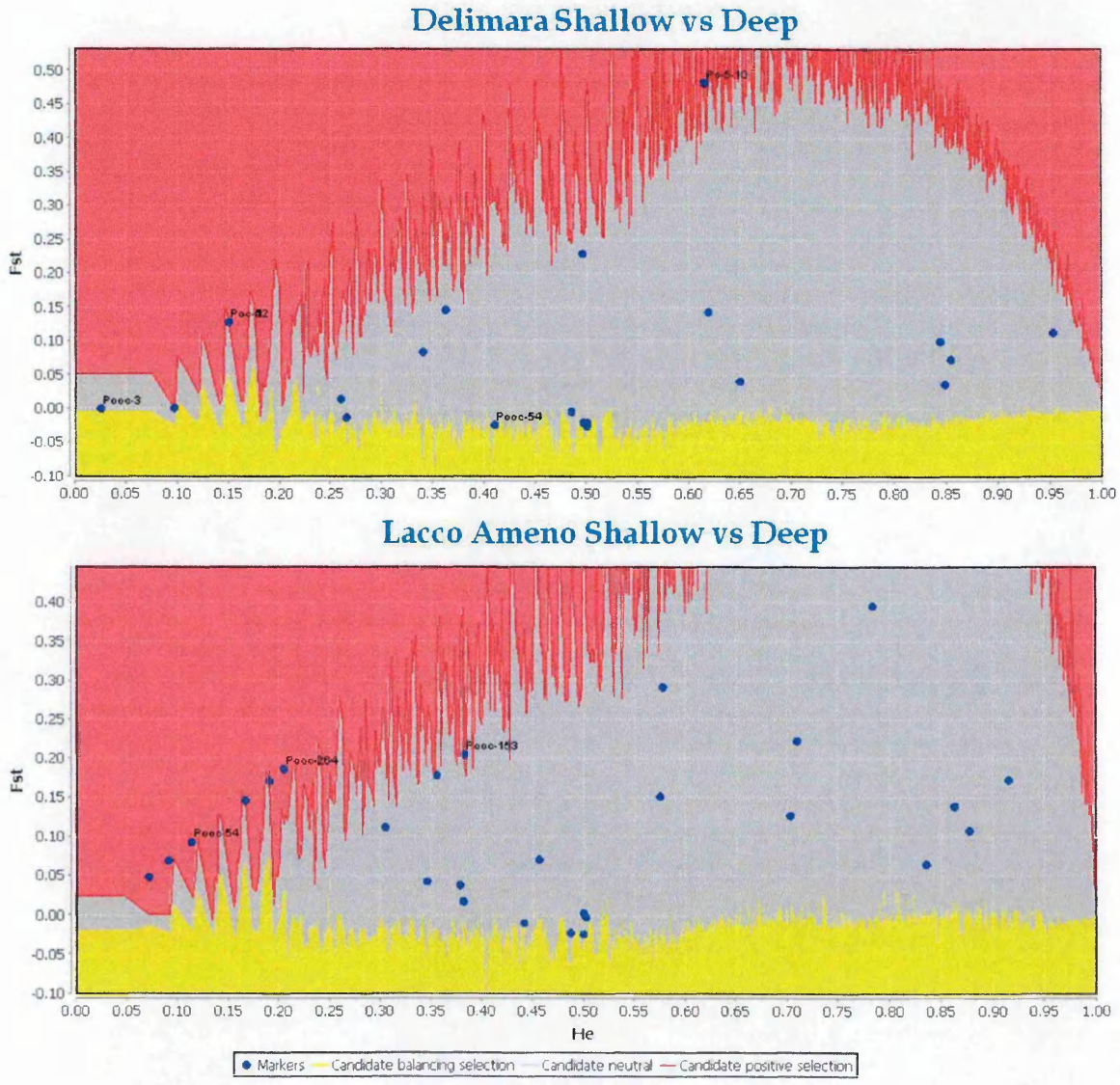
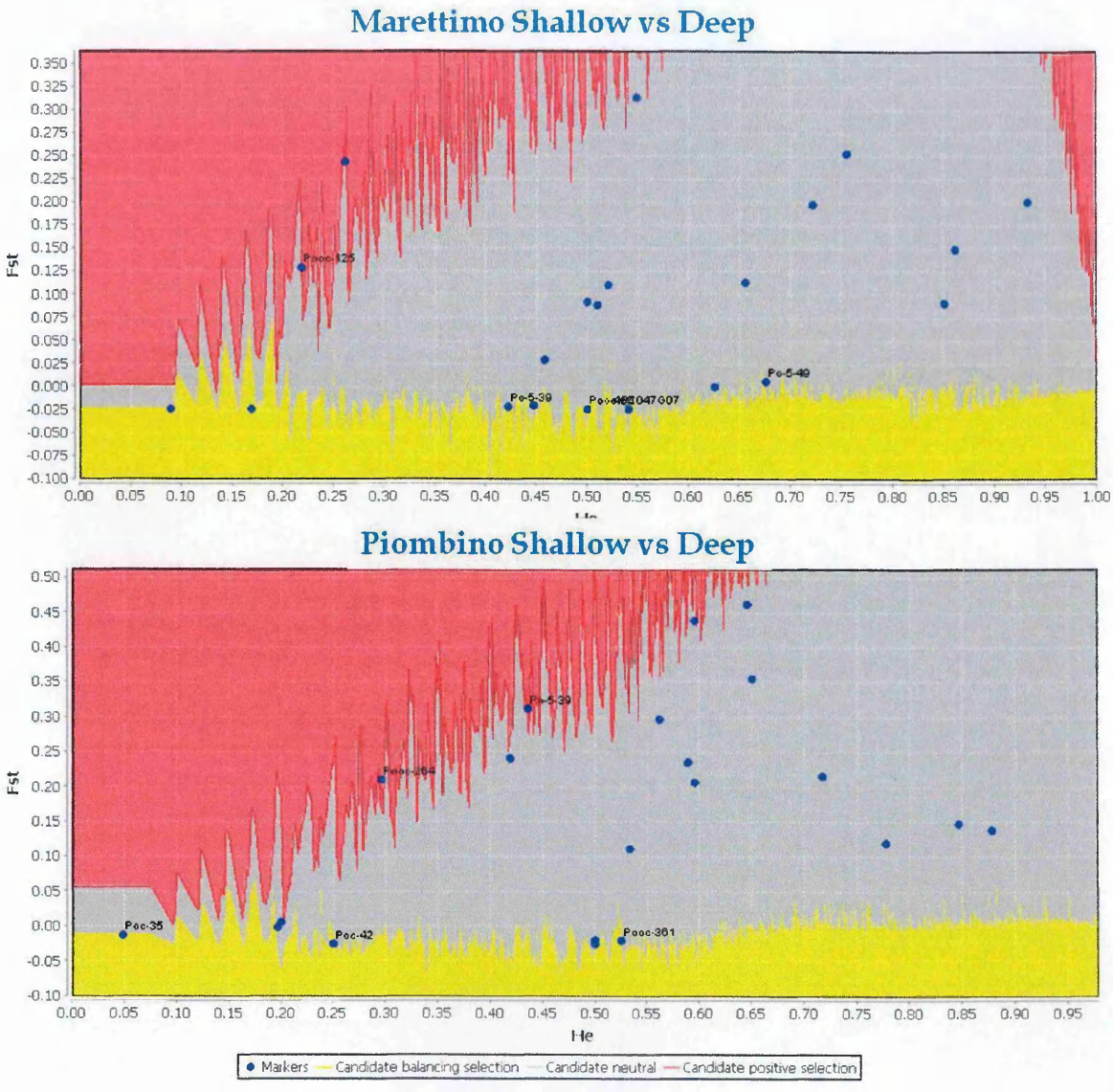
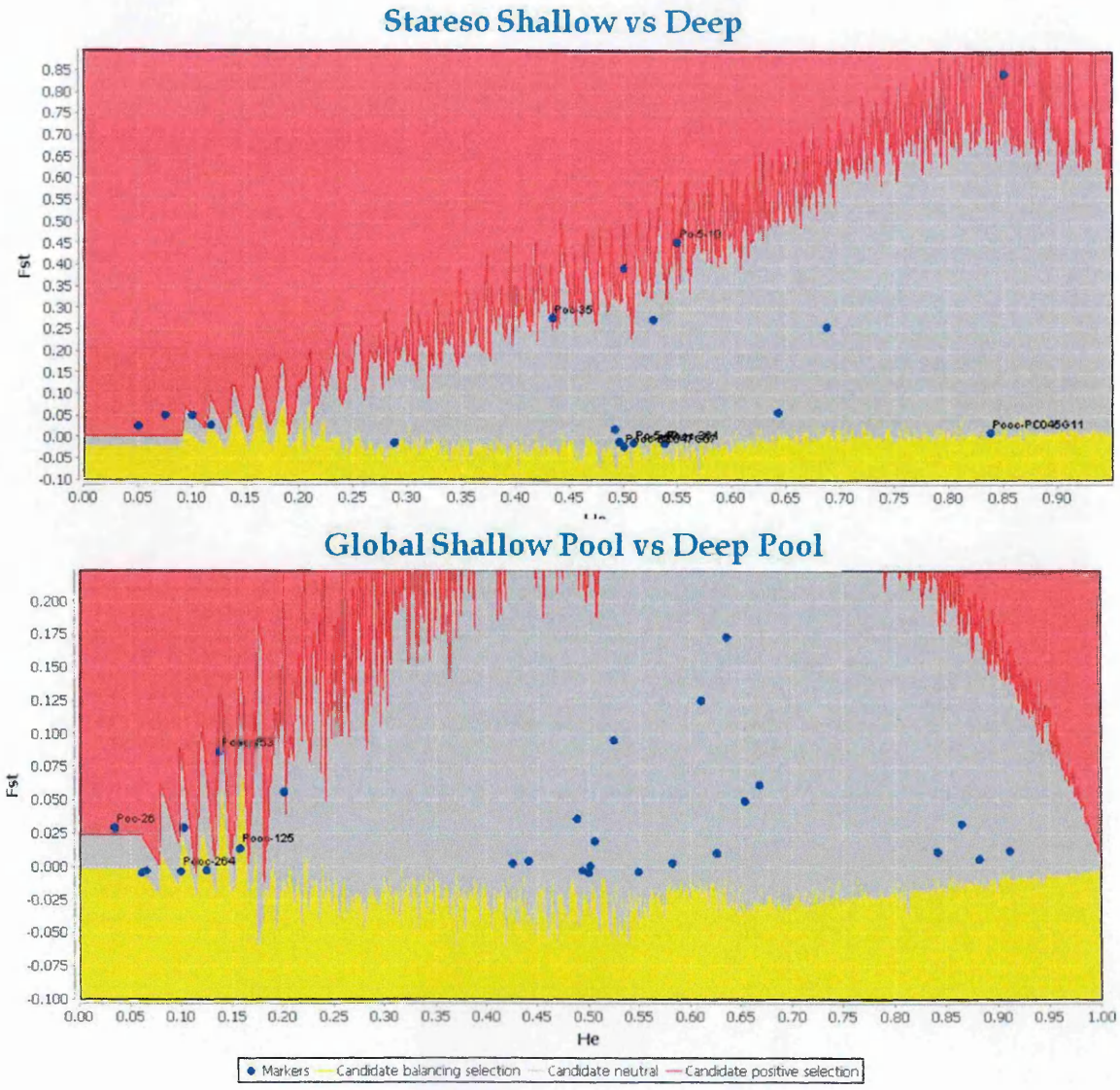


Figure 4.A1. Visual representation of the screening for microsatellites polymorphism in the 6 populations and at the shallow and deep site of each population. Full circles = polymorphic loci; empty circle = monomorphic loci; nP = number of polymorphic loci; %P = percentage of polymorphic loci; S = shallow site; D = deep site. A-SSR = set of anonymous microsatellites; E-SSR = set of EST-linked microsatellites.

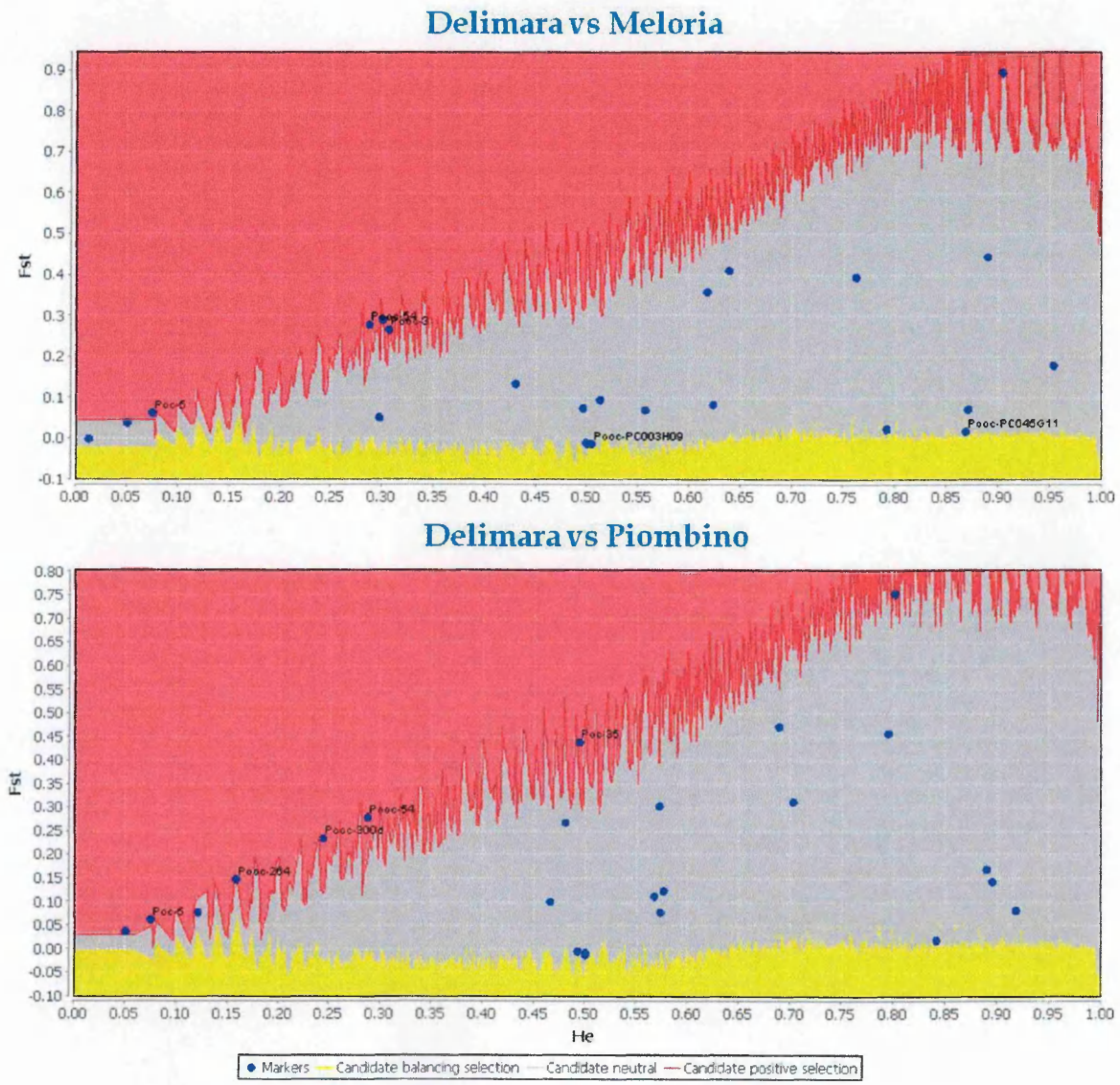
Plot 4.A1. Outlier detection with LOSITAN in 6 pair-wise-comparisons along the bathymetric gradient. Distribution of empirical F_{ST} values is shown as function of expected heterozygosity. For each comparison the result of one of the three runs at 99% CI is reported. Each dot indicates a microsatellite marker. The gray shaded area represents the simulated neutral distribution, whereas loci falling in the red and yellow shaded area are candidates for selection.

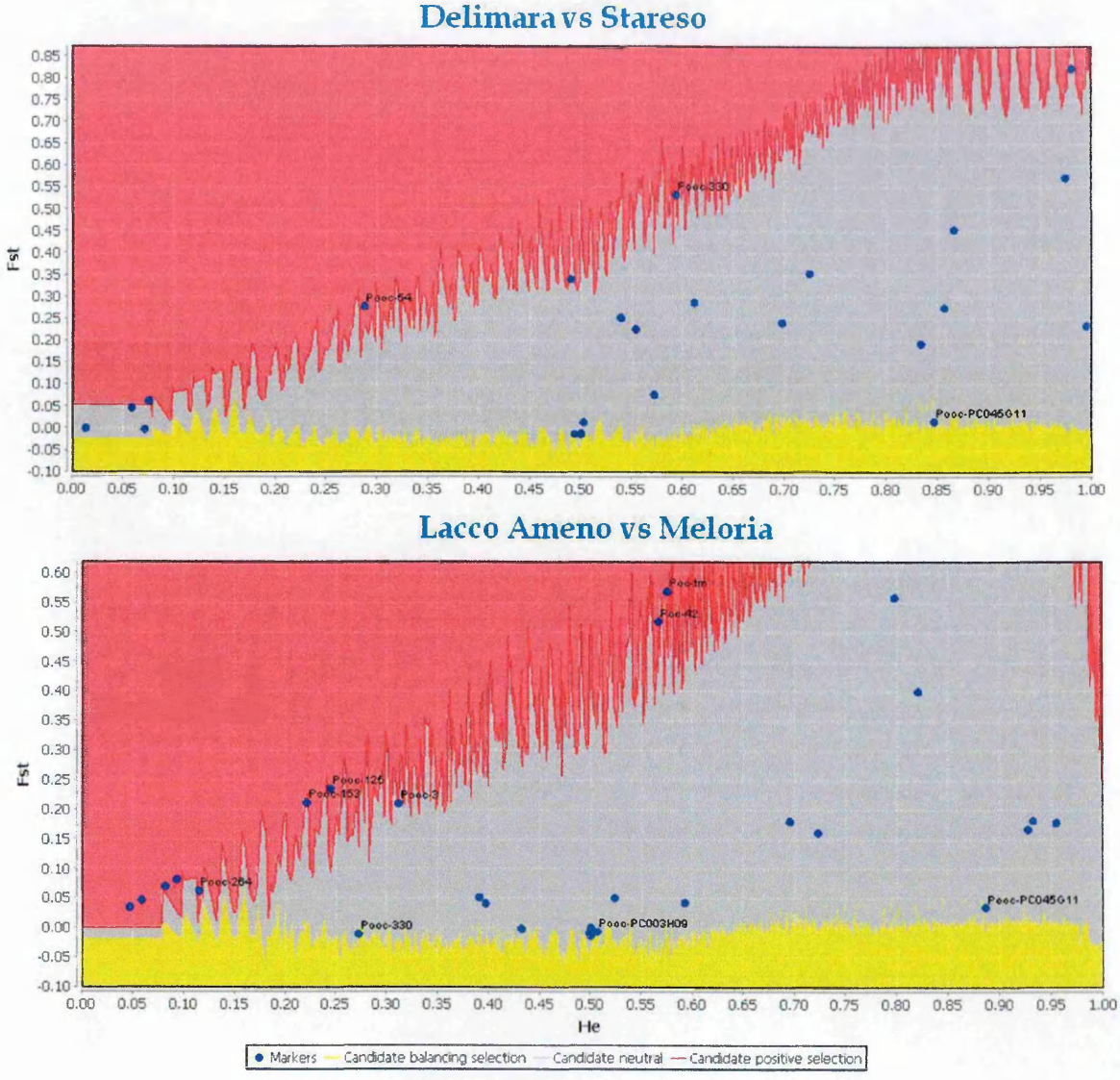


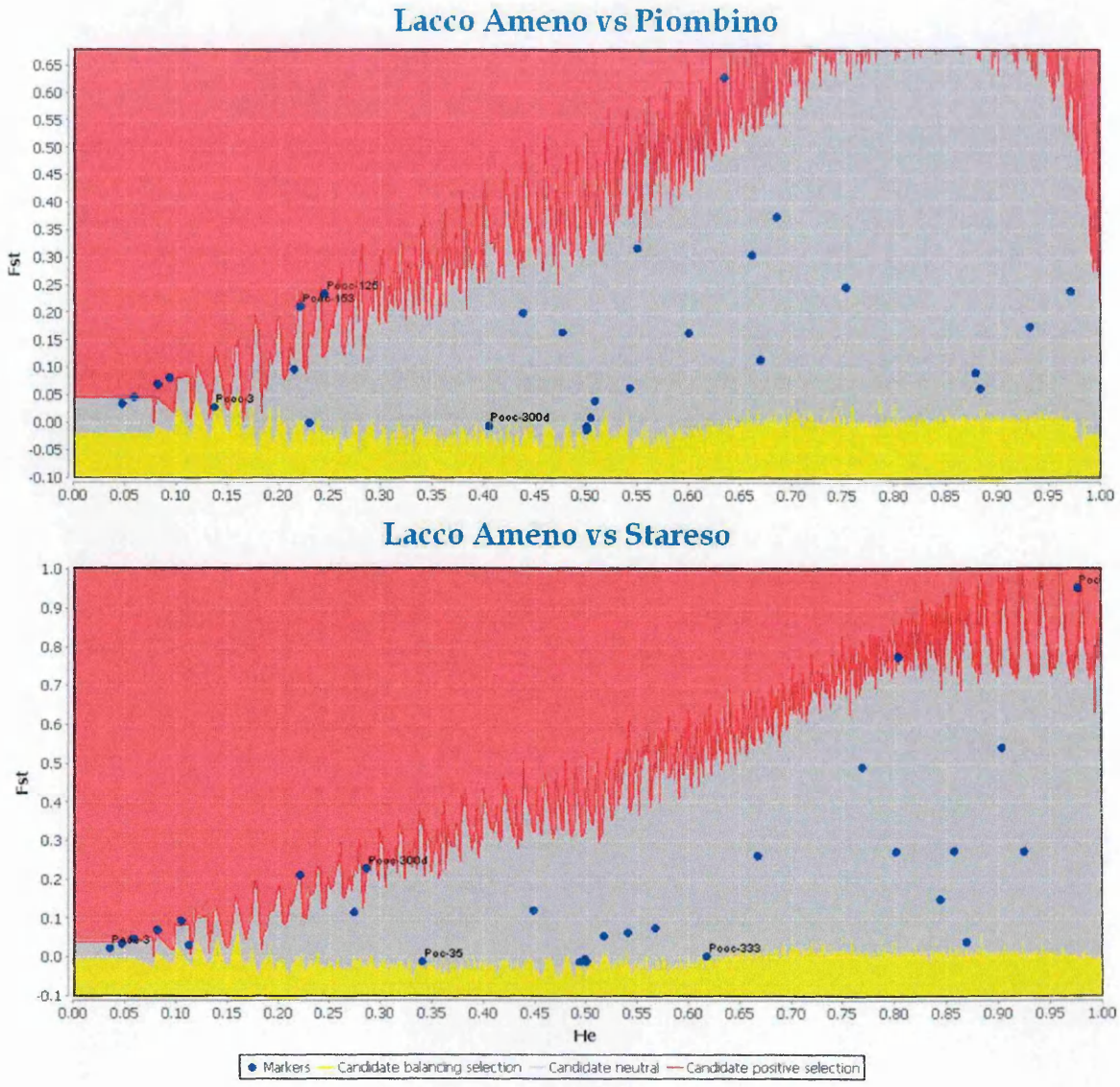

Marettimo Shallow vs Deep
Fst
He
Pooc-125
Po-5-39
Po-5-49
Piombino Shallow vs Deep
Fst
He
Po-5-39
Poc-35
Poc-42
Pooc-361
Markers
Candidate balancing selection
Candidate neutral
Candidate positive selection

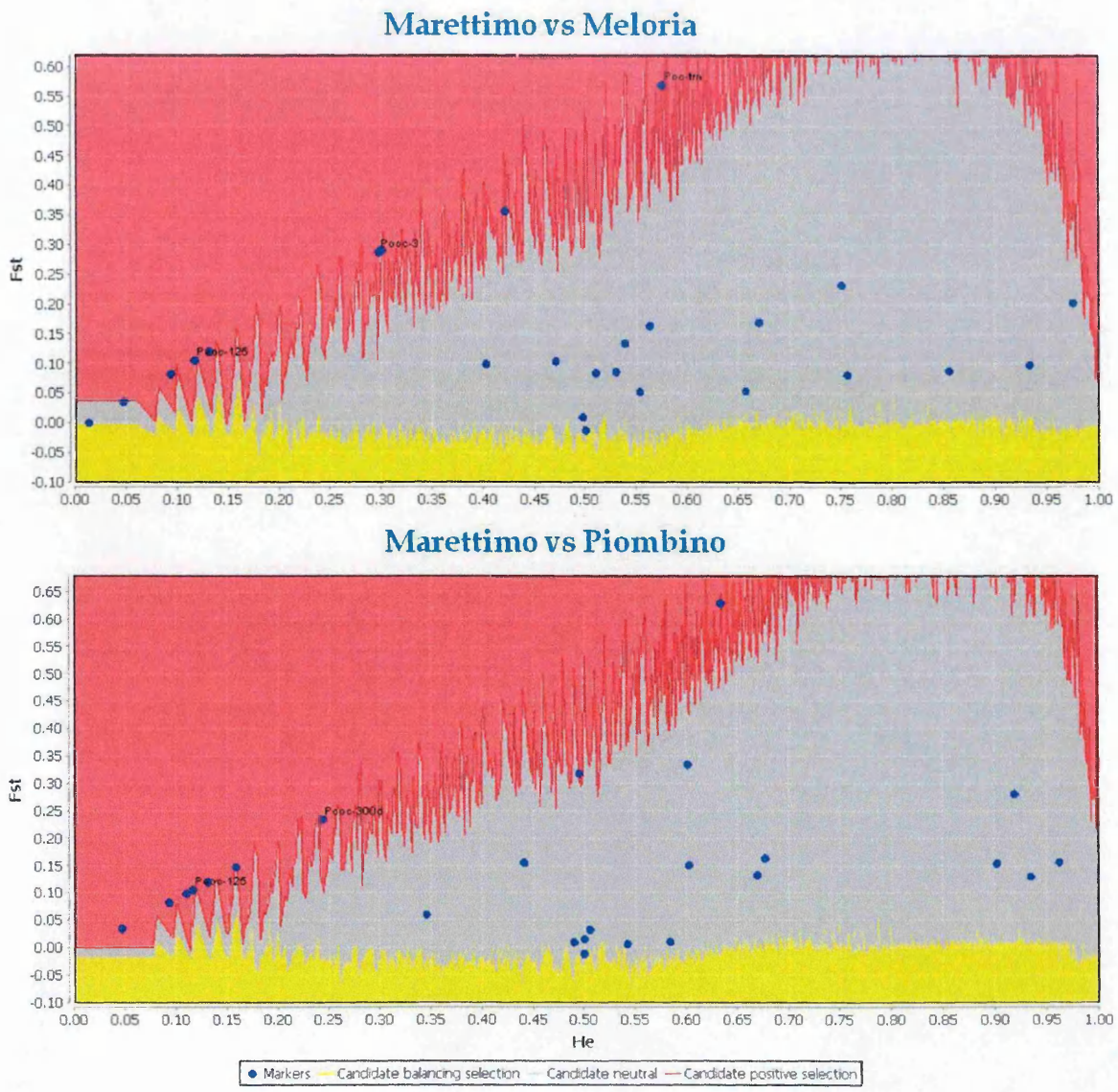

Stareso Shallow vs Deep
Fst
He
Poc-35
Poc-PC045G11
Global Shallow Pool vs Deep Pool
Fst
He
Poc-26
Poc-125
Poc-264
Markers
Candidate balancing selection
Candidate neutral
Candidate positive selection

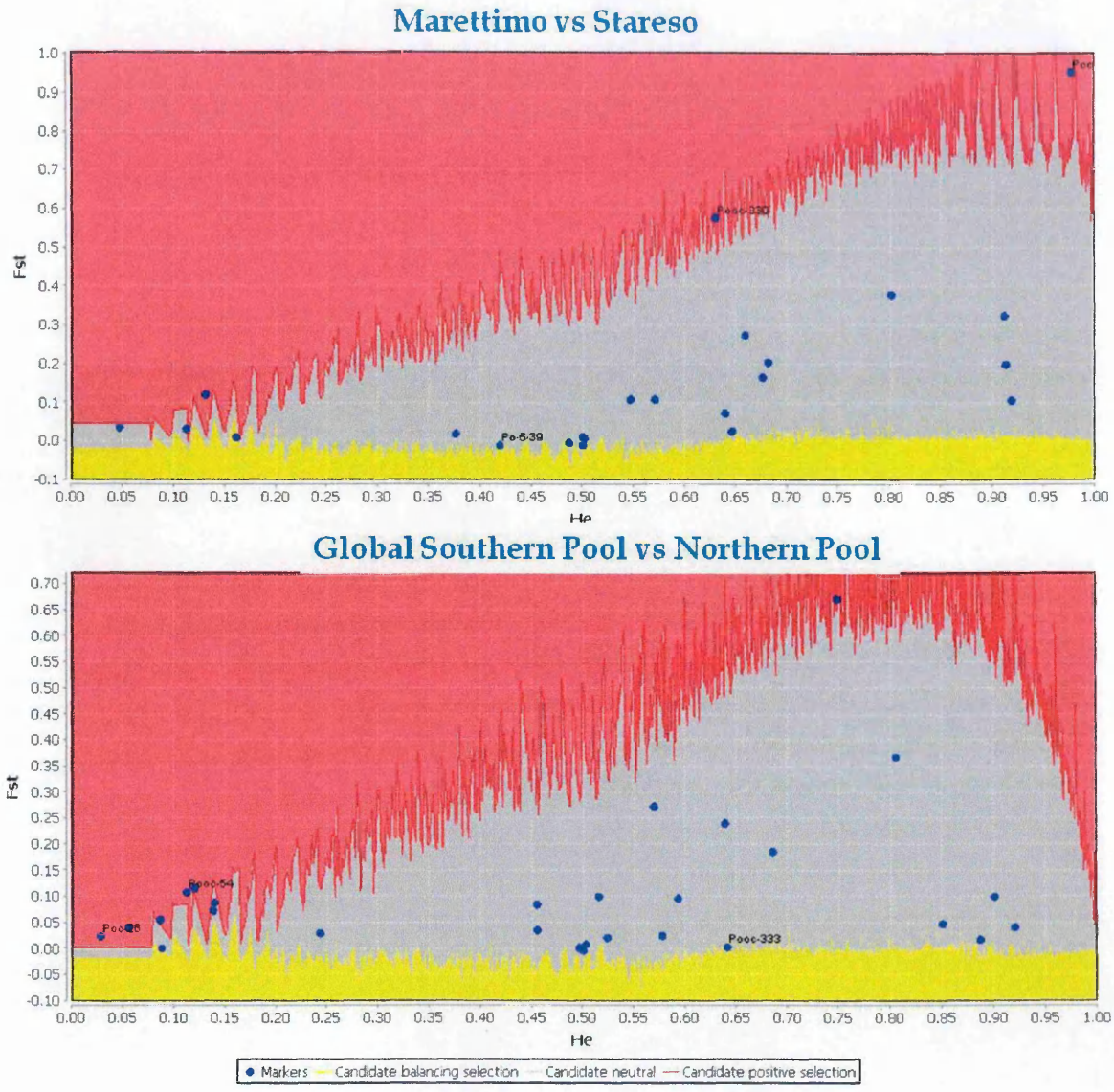
Plot 4.A2. Outlier detection with LOSITAN in 10 pair-wise-comparisons along the latitudinal gradient. One of the three runs is showed for each comparisons. Distribution of empirical F_{ST} values is shown as function of expected heterozygosity. For each comparison the result of three runs at 99% CI is reported. Each dot indicates a microsatellite marker. The gray shaded area represents the simulated neutral distribution, whereas loci falling in the red and yellow shaded area are candidates for selection.

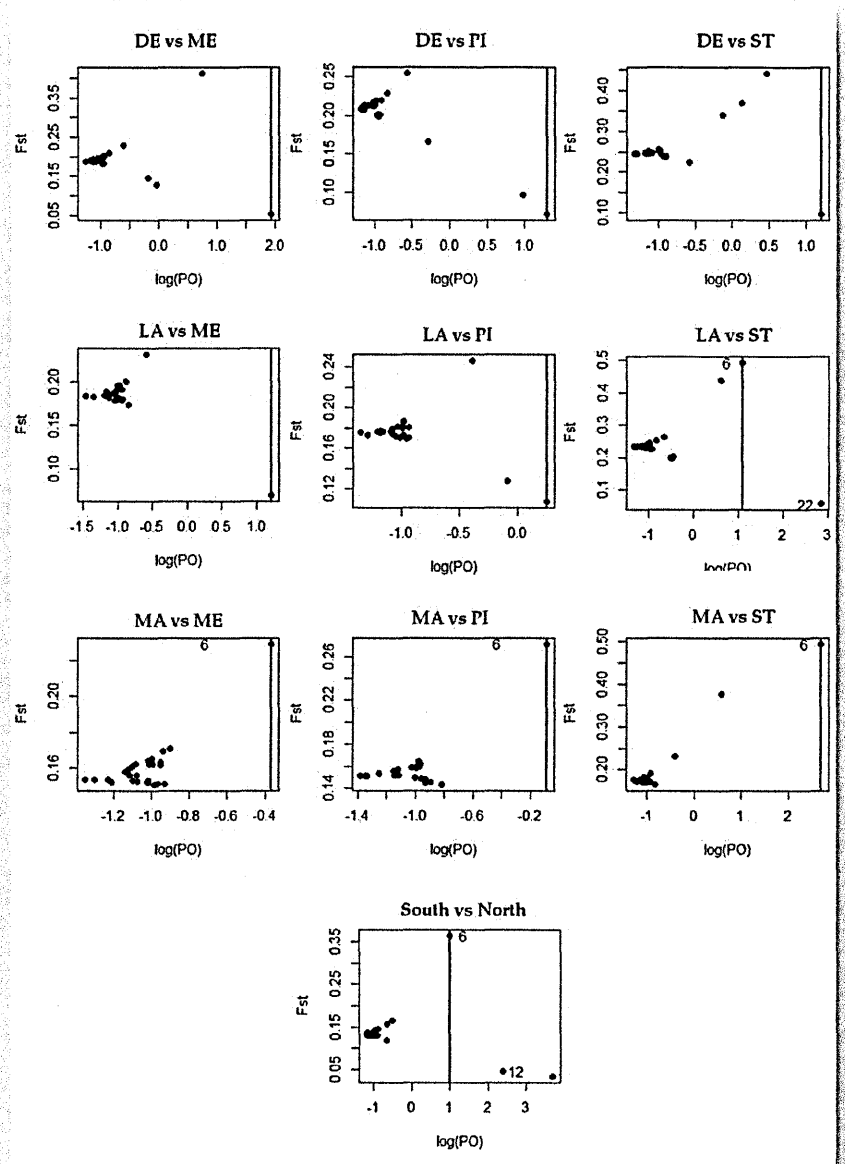



Delimara vs Stareso
Fst
Pooc-330
Pooc-54
Pooc-PC045G11
Lacco Ameno vs Meloria
Fst
He
Pooc-tm
Pooc-42
Pooc-125
Pooc-152
Pooc-3
Pooc-284
Pooc-330
Pooc-PC003H09
Pooc-PC045G11
Markers
Candidate balancing selection
Candidate neutral
Candidate positive selection


Lacco Ameno vs Piombino
Fst
He
Poec-3
Poec-300d
Poec-35
Poec-333
Markers
Candidate balancing selection
Candidate neutral
Candidate positive selection


Marettimo vs Meloria
Fst
He
Poec-3
Poec-125
Marettimo vs Piombino
Fst
He
Poec-3006
Poec-125
Markers
Candidate balancing selection
Candidate neutral
Candidate positive selection


Marettimo vs Stareso
Fst
He
Poec-330
Po-5-39
Global Southern Pool vs Northern Pool
Fst
He
Poec-54
Poec-333
Markers
Candidate balancing selection
Candidate neutral
Candidate positive selection



Plot 4.A3. BAYESCAN results for the 10 pair-wise comparisons. One of the three runs is showed for each comparison. F_{ST} = locus-specific genetic divergence among populations; log(PO) = decision factor in logarithmic scale (base 10) to determine selection; a vertical line indicates evidence for selection. Filled circles represent loci analyzed. DE = Delimara, LA = Lacco Ameno, MA = Marettimo, ME = Meloria, PI = Piombino, ST = Stareso..

Table 4.A1 Standardized lnRH, lnRV and lnRθ values for the pair-wise comparisons involved in the genome scan along the latitudinal gradient. For completeness all the possible comparisons are reported. For information about the comparisons refer to Plot 4.A3 legend.

	DE vs. ME			DE vs. PI			DE vs. ST			LA vs. ME			LA vs. PI		
	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ
Locus															
***Poc*-35**															
Po5-49															
Poc-trn										-2.766	-3.115	-2.835	-2.873	-3.059	-2.967
Pooc-264				-1.975	-2.470										
Pooc-330															
Pooc-3			-2.362												
Pooc-54	-2.556		2.209	2.293		2.068	2.845		2.041						
Pooc-153	2.476										2.491		1.973	2.553	
Pooc-125							-1.968	-3.229			2.522		2.024	2.584	
Pooc-300	-2.534	-3.456	-2.351	-2.161	-3.174	-2.046									
Pooc-361															

continued

	LA vs. ST			MA vs. ME			MA vs. PI			MA vs. ST			Global S vs. N		
	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ
Locus															
Poc-35									2.151						
Po-5-49															
Poc-trn	-2.023	-2.328		-2.320	-2.513	-2.375	-2.178	-2.347	-2.332	-2.307	-2.701	-2.060			
Pooc-264															
Pooc-330										1.990					
Pooc-3				-2.213		-2.155									-2.020
Pooc-54													2.327		2.009
Pooc-153	2.133	2.370			2.226			2.178		2.688	2.969		2.372	3.807	2.067
Pooc-125															
Pooc-300d	2.324	2.073	2.261	-2.215	-2.446	-2.160	-2.029	-2.254	-2.038			2.004			
Pooc-361		-2.012													

Table 4.A2. Standardized lnRH, lnRV and lnRθ values for the pair-wise comparisons involved in the genome scan along the depth gradient. S = Shallow, D = deep.

	Delimara S vs. D			Lacco Ameno S vs. D			Marettimo S vs. D			Piombino S vs. D			Stareso S vs. D			Global S vs.D		
	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ	lnRH	lnRV	lnRθ
Locus																		
***Poc*-35**													2,245	2,085	2,237			
Po5-49												1,975						
***Poc*-26**																3,073	2,905	2,510
Po15														2,135				
Po5-10	-2,819	-2,071	-2,466					2,697			-1,991		2,214		2,130			
Po4-3				2,664														
Po5-39															-2,359			
Po5-40								1,999										
Pooc-264						-1,991				2,817								
Pooc-330				-2,208														
Pooc-153				-2,477					-2,609							-3,298	-3,048	-3,218
Pooc-333		-2,353																
Pooc-125							-4,183	-2,354	-2,280									
Pooc-214	2,269					2,006												
Pooc-300					-2,108													
Pooc-PC047G07					-3,234													

Table 4.A3. Outlier loci detected in the 6 pair-wise comparisons related to depth. In the first column the pair-wise comparisons are reported: DE = Delimara shallow vs. deep; LA = Lacco Ameno shallow vs. deep; MA = Marettimo shallow vs. deep, ME = Meloria shallow vs. deep, PI = Piombino shallow vs. deep; ST = Stareso shallow

vs. deep; G = shallow pool vs. deep pool. In the second column the three approaches applied for outliers detection are indicated: LS = Beaumont and Nichols (1996) method implemented in the software LOSITAN; RH, RV, Rθ = Sclötterer tests lnRH, lnRV and LnRθ; BS = Foll and Gaggiotti (2008) method implemented in the software BAYESCAN. EST-derived loci are in italics. **=95% level of significance, ***=99% level of significance, ****=99.9% level of significance.

		Genome scan along a depth gradient																														
		Poc-45	Po-5	Poc-5	Poc-35	Po-5-49	Poc-tm	Poc-26	Po-15	Po-5-10	Po-4-3	Po-5-39	Po-5-40	Poc-42	*Pooc-229*	*Pooc-264*	*Pooc-330*	*Pooc-3*	*Pooc-54*	*Pooc-153*	*Pooc-333*	*Pooc-125*	*Pooc-PC045G11*	*Pooc-PC044B02*	*Pooc-214*	*Pooc-300*	*Pooc-PC047G07*	*Pooc-50*	*Pooc-361*	*Pooc-PC003H09*		
PC	M																															
DE	LS	**			**					***								***	***									**				
	RH			**						**				**																		
	RV													***																		
	Rθ									**																						
	BS																															
LA	LS																		***	***												
	RH							**								**									**							
	RV															**										**	***					
	Rθ															**									**							
	BS																															
MA	LS	***																				***					***	***				
	RH																			***		***										
	RV																			****												
	Rθ																			***		**										
	BS																															
PI	LS	**			***							***		***		***									**			**	**	***		
	RH															***																
	RV									**																						
	Rθ					**																										
	BS																															

ST	LS		***				**				***
	RH				**	**					
	RV										
	Rθ	**			**	**					
	BS						*				
G	LS			***				***		***	
	RH			***					***		
	RV			***					***		
	Rθ			**					***		
	BS										

Table 4.A4. Outlier loci detected in the 10 pair-wise comparisons related to latitude. In the first column the pair-wise comparisons (PC) are reported: DE-ME, DE-PI, DE-ST = comparisons involving Delimara vs. Meloria, Piombino and Stareso, respectively. LA-ME, LA-PI, LA-ST = comparisons involving Lacco Ameno vs. Meloria, Piombino and Stareso, respectively. MA-ME, MA-PI, MA-ST = comparisons involving Marettimo vs. Meloria, Piombino and Stareso, respectively, S-N = comparison involving the Southern and the Northern populations pools. In the second column the three approaches (M) used for outliers detection are indicated: LS = Beaumont and Nichols method (1996) implemented in LOSITAN; BS = Foll and Gaggiotti method (2008) implemented in BAYESCAN; RH, RV, Rθ = Sclottered tests lnRH, lnRV and lnRθ. EST-derived loci are in italics. *=90% level of significance,**=95% level of significance, ***=99% level of significance, ****=99.9% level of significance.

		Poc-45	Po-5	Poc-5	Poc-35	Po-5-49	Poc-trn	Poc-26	Po-15	Po-5-10	Po-4-3	Po-5-39	Po-5-40	Poc-42	*Pooc-229*	*Pooc-264*	*Pooc-330*	*Pooc-3*	*Pooc-54*	*Pooc-153*	*Pooc-333*	*Pooc-125*	*Pooc-PC045G11*	*Pooc-PC044B02*	*Pooc-214*	*Pooc-300*	*Pooc-PC047G07*	*Pooc-50*	*Pooc-361*	*Pooc-PC003H09*
PC	M																													
DE-ME	LS																	***	***				***							**
	RH																	**	**							**				
	RV																									****				
	Rθ																	**	**							**				
	BS																						***							
DE-PI	LS			***		***													***							***				
	RH															**			**							**				
	RV															**										***				
	Rθ																		**							**				
	BS																						*							
DE-ST	LS			**															***				***							
	RH																		***			**								
	RV																					**								

	Rθ										**							
	BS														*			
LA-ME	LS		***				***	***	***	**		***		***	***			**
	RH		***															
	RV		***									**		**				
	Rθ		**															
	BS														*			
LA-PI	LS		**							***		***		***				
	RH		***									**		**				
	RV		***									**		**				
	Rθ		***															
	BS																	
LA-ST	LS	***	***							***			**			**		
	RH		**									**				**		
	RV		**									**				**	**	
	Rθ															**		
	BS		*												***			
MA-ME	LS		***							***				**				
	RH		**							**						**		
	RV		**									**				**		
	Rθ		**							**						**		
	BS																	
MA-PI	LS	**	**											***		**		
	RH		**													**		
	RV		**									**				***		
	Rθ	**	**													**		
	BS																	
MA-ST	LS		***		***				***									
	RH		**						**			***						
	RV		***									***						
	Rθ		**						**									
	BS		***															
S-N	LS		**	***									***					
	RH										**	**						
	RV											***						
	Rθ									**	**	**						
	BS		*			***									***			

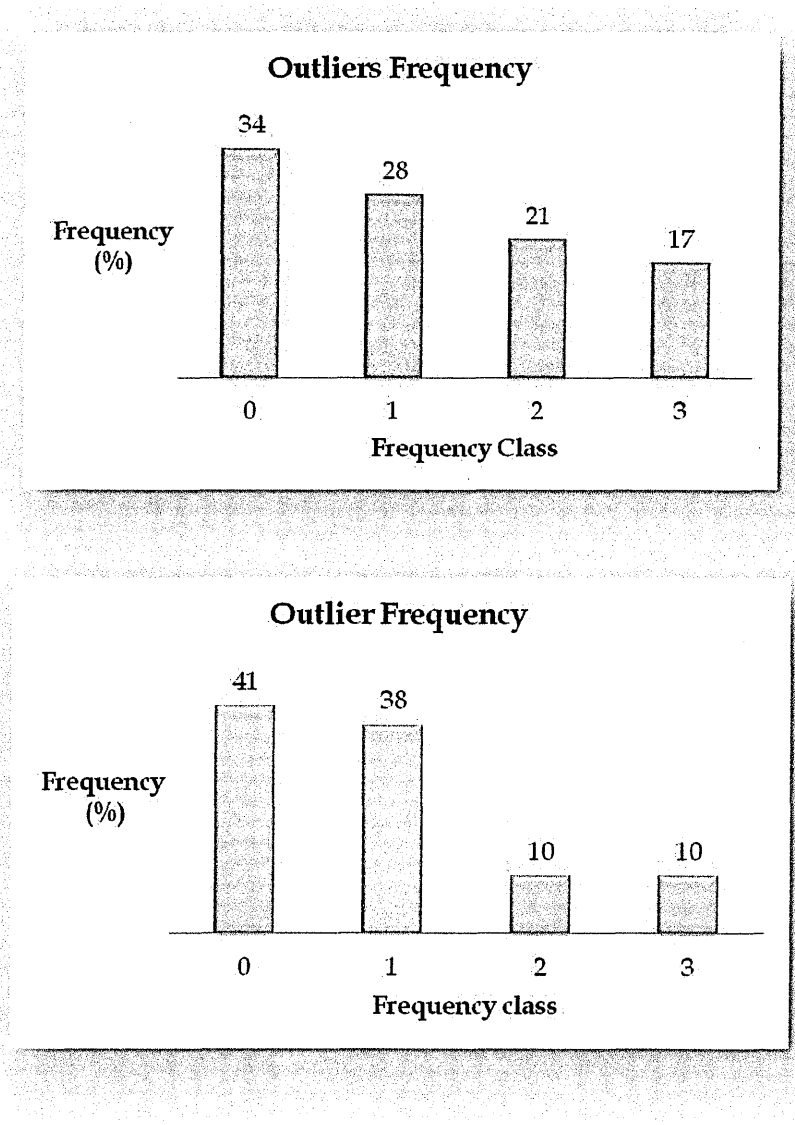
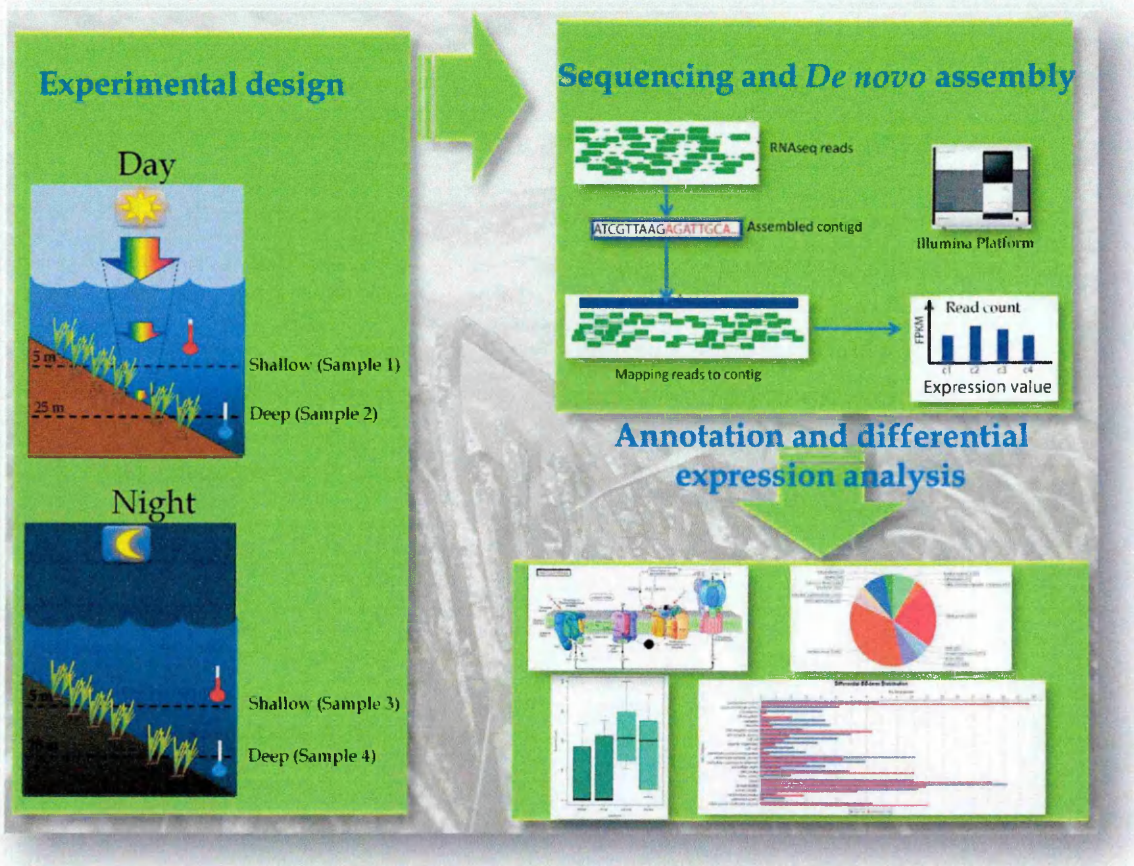


Figure 4.A2. Outliers frequency distribution from the genome scan along the bathymetric (upper panel) and latitudinal gradient (lower panel). The frequency (%) of classes of outliers were detected on the basis of the number of comparisons in which they were present.

Chapter 5


Experimental design
Day
5 m
25 m
Shallow (Sample 1)
Deep (Sample 2)
Night
Shallow (Sample 3)
Deep (Sample 4)
Sequencing and De novo assembly
RNAseq reads
ATCGTTAAGAGATTGCA..
Assembled contigd
Illumina Platform
Mapping reads to contig
Read count
FPKM
Expression value
Annotation and differential expression analysis

Transcriptome sequencing and analysis of *Posidonia oceanica*. Insights into environmental adaptation.

5.1-Abstract

The development of massively parallel sequencing technologies enables the sequencing of total cDNA (RNA-Seq) to derive accurate measure of individual gene expression. Here, the first use of RNA-Seq in *Posidonia oceanica* is reported. The goals of the analysis were to increase genomic resources in this species and to gain insight into the wide range of transcriptional responses in plants that were collected at two times during the day and at two different depths and were associated with different conditions of light and temperatures, two key environmental factors that govern plant growth and development. More than 207 million sequencing reads were generated using Illumina paired-end sequencing technology. In total, 53001 contigs were obtained. Of these, 18485 (34.8%) were successfully annotated. Gene Ontology (GO) terms enrichment analysis highlighted the capability of plants collected at shallow depth to cope with environmental stresses imposed by high light and high temperature by a transcriptional remodeling aimed to trigger signaling cascades to ensure their survival. A shift was revealed in relation to the time of the day to initiate the cell cycle, with shallow plants more active in cell cycling during the day and deep plants at night.

Keywords: de novo transcriptome; light; temperature; depth; enrichment analysis.

5.2-Introduction

Plant growth and development respond tightly to the surrounding environment. This is achieved through the complex integration of multiple environmental stimuli, of which light and temperature are arguably the most important. Light signals provide plants with spatial, temporal and seasonal information (Franklin and Whitelam, 2003). Plants measure the quantity, quality, direction and photoperiod of light and use this information to regulate multiple developmental processes throughout their life-cycle (Whitelam and Halliday, 2007). Light, additionally, provides entrainment signals to the circadian clock, ensuring coordination of growth with light/dark cycles and ultimately enhancing plant fitness (Dodd et al., 2005). Temperature is an important seasonal cue which enables plants to predict and consequently prevent the adverse effects of environmental change. Light and temperature regulate similar developmental processes throughout the life-cycle of plants due to complex crosstalk between signaling pathways, which optimize plant development in natural environments (Franklin, 2009). For this reason it is not always possible to separate the influence of one factor from the other but the effects have to be analyzed synergistically.

This is true also for marine plants. Seagrass growth, in fact, depends on the quantity and quality of light available for photosynthesis (Zieman and Wetzel, 1980). Reduction of underwater light often results in a large-scale seagrass die-off (Short and Wyllie-Echeverria, 1996). Underwater light intensity is attenuated exponentially with water depth, and seagrasses grow along the bathymetric

gradient, from the intertidal zone to depths at which plants receive minimum irradiance (Lee et al., 2007). Minimum light requirements for these plants are much higher (2-37% surface irradiance, SI) than those of macroalgae and phytoplankton (1-3%, SI; Strickland, 1958; Luning and Dring, 1979; Sand-Jensen, 1988; Duarte, 1991; Markager and Sand-Jensen, 1992; Dennison et al., 1993).

Seagrasses can enhance light harvesting efficiencies through photo-acclimation during low light conditions (West, 1990; Olesen and Sand- Jensen, 1993; Olesen et al., 2002). Plants collected near their maximum depth limits have higher photosynthetic efficiencies (α) and lower light requirements for saturate photosynthesis than plants growing in shallower waters (Masini and Manning, 1997). Lower compensation irradiance (I_c) and greater photosynthetic efficiencies have also been observed in seagrasses inhabiting inner harbors, where light availability is reduced through high turbidities, in comparison to less turbid outer harbor stations (Ruiz and Romero, 2003).

Growth rates and photosynthesis in seagrasses change across seasons and depths, due to the interaction between light and temperature. Many species exhibit clear seasonal trends, with photosynthetic parameters that increase during the warmer months (Dennison, 1987; Herzka and Dunton, 1997; Alcoverro et al., 1998), and with growth rate increasing during spring and summer and decreasing growth during fall and winter (Orth and Moore, 1986; Vermaat et al., 1987; Macauley et al., 1988; Dunton, 1990; Thom, 1990; Lee and Dunton, 1996). Since the effects of irradiance and temperature are difficult to disentangle, based on studies involving water temperature with seasonal growth patterns numerous researchers consider temperature a primary factor controlling seasonal growth (Setchell, 1929; Tutin,

1942; Phillips et al., 1983; Bulthuis, 1987; Lee and Dunton, 1996). Both respiration and photosynthesis increase with increasing water temperatures, but respiration usually increases more than photosynthesis at progressively higher temperatures, thus leading to a reduction in net photosynthesis (Bulthuis, 1983; Dennison, 1987; Marsh et al., 1986; Pérez and Romero, 1992; Herzka and Dunton, 1997; Masini and Manning, 1997). From the available data, the optimum temperatures for both growth and photosynthesis are variable with seagrass species, but the optimum temperatures for growth were usually lower than those for photosynthesis, and for *P. oceanica* optimum growth is reported at 15.5 ± 2.5°C (Lee et al 2007).

Photosynthetic parameters, such as light-saturated photosynthetic rate (P_{max}), saturation irradiance for photosynthesis (I_k), compensation irradiance (I_c) and respiration rates usually increase with increasing water temperatures (Bulthuis, 1983; Dennison, 1987; Marsh et al., 1986; Pérez and Romero, 1992; Herzka and Dunton, 1997; Masini and Manning, 1997; Moore et al., 1997), although the optimum temperature for photosynthesis can vary with underwater irradiance (Bulthuis, 1987). Seagrasses growing in low light conditions have lower optimum temperatures for photosynthesis than plants in high light conditions. Plants at higher temperatures likely need more light to maintain positive carbon balance than those at lower temperatures. Thus photosynthetic production in seagrasses is more susceptible to high water temperatures at reduced light conditions (Lee et al., 2007). Consequently, reductions in underwater light may be more harmful to seagrasses during summer rather than in winter (Hillman et al., 1989).

In general, the role of temperature on seagrass growth is considerably more complicated than its effect on photosynthesis. The optimal temperature for

seagrasses results from the effect of temperature on nutrient availability and uptake, leaf senescence, nutrient partitioning within the plants, and respiration (Marsh et al., 1986; Bulthuis, 1987; Herzka and Dunton, 1997).

Because seagrasses have a sessile lifestyle and cannot easily disperse from a given locality, they must constantly adapt their development, growth and architecture to survive an every changing environment. The remodeling of the transcriptome by appropriate changes in gene expression is a way that plants use to respond and adapt to environmental stimuli. The activation of cascades of molecular networks involved in stimuli perception, signal transduction and the expression of specific genes and metabolites allow plants to mediate the response to the stimulus (Dos Reis et al., 2012). Gene expression studies are proving valuable in examining plastic and adaptive response to global climatic changes as showed in *Zostera marina* with respect to temperature stress (Franssen et al. 2011) and in *Posidonia oceanica* with respect to depth and light availability (Procaccini et al. 2010).

In *Posidonia oceanica,* a first step towards the comprehension of the environmental adaptation to light and temperature at molecular level have been provided by a single EST library constructed by pooling RNAs from plants living at different depths. A total of 5185 raw sequences were obtained and assembled in 1219 tentative unigenes, which have been annotated and are now available in the seagrass sequence database Dr. Zompo (Wissler et al., 2009). In addition, the construction of a SSH (Suppression Subtractive Hybridization) cDNA library, using samples collected at two different depths (-5m and -25m) from a continuous meadow (Procaccini et al., 2010; Dattolo et al., 2013 *in press*), showed that different light and temperature regimes leaded to important changes in plant gene

expression profiles. Out of 486 contigs produced from the SSH library, only 28 (5.86%) were common to both shallow and deep libraries. The highest differences between the two depths were in primary metabolism, photosynthesis and stress defense genes. In addition, this study generated 850 TUGs (Tentative Unigenes) that were also included in the Dr. Zompo database (Dattolo et al., 2013 *in press).* The existing EST libraries provided the starting point for further studies aimed to investigate specific processes that can be involved in the response to light and temperature. For example, gene expression analysis of genes involved in different phases of the photosynthetic process, carried out by real time PCR, showed overall an up-regulation in high light, during the whole daily cycle (Ruocco et al., 2012). Nevertheless, the complete understanding of the role of environmental stress and the intrinsic capacity of seagrasses to respond to climate change requires an integrative and interdisciplinary research. Currently, seagrass research community aims to integrate the fields of seagrass ecophysiology and ecogenomics in order to address questions related to anthropogenic disturbances that alter light, temperature and pH, thereby affecting primary production as well as permanent stress (selection) regimes that may exceed the capacity of seagrass to respond (and eventually adapt) while still providing ecological services (Procaccini et al., 2012). For example transcriptomic approaches may be paralleled by a proteomic one to bring genetic and physiological processes closer together, as in the case of *P. oceanica* for protein analysis under chronic low light conditions (Mazzuca et al. 2009).

Information on the transcriptome of this species is still scarce, with a total of 2069 genes annotated and available in the Dr. Zompo database (Wissler et al., 2009).

Next generation RNA sequencing (NGS, RNA-Seq) technologies, are recently become methods of choice for gene expression analysis due to their unprecedented level of sensitivity and high-throughput nature (Jain, 2012). RNA-seq has also been shown to be a very efficient, cost-effective approach for transcriptome analysis of non-model species. In fact, also in absence of the genome, RNA-seq allows construction of complete transcriptome of an organism by *de novo* assembly (Martin and Wang, 2011). Since variation in the external environment can alter the transcriptome of a cell, transcriptome studies can also provide a snapshot of the organism's response, based on the genes that are expressed in conditions and times of interest. Although there is not always a linear relationship between the expressed genes and the proteins synthesized, such studies give us a reasonable insight into the state of the cell allowing to identify which are the main pathways involved in the response to the plants at different conditions of light and temperature.

In this study, a complete transcriptome sequencing of *Posidonia oceanica* plants living at two different depths and in different times of the day was performed applying an Illumina© massive parallel sequencing technology. Since different samples collected in different times and conditions were utilized for the study, our analysis also allowed for a comparison among genes and pathways involved in plant metabolism in different light and temperature conditions.

The objectives of this study were to increase the genomic resources for *P. oceanica* and to explore the transcriptomic response to different conditions of light and temperature. Infact, regulatory changes in the gene expression, in addition to structural changes in genes, are known to be involved in the adaptation to

environmental changes. This is particularly true for plants that, being sessile organisms, have to provide a rapid response to environmental changes that is typically mediated at gene expression level. Evidence for a regulation in gene expression at light and temperature changes was already provided in *P. oceanica*, as stated above, from real time experiments and from the construction SSH library. Thus, here we expected to find different expression profiles between plants experiencing different conditions of light and temperature and also we expected to provide a more deep view of the processes involved..

5.3-Materials and Methods

5.3.1-Sampling design and RNA extraction

Posidonia oceanica shoots were collected in October 2012 from a meadow located in front of the Station de Recherches Sous-marine et Océanographiques (STARESO; Corse, 8°45'E, 42°35'N, Figure 5.1). Sampling was performed at two different depths, -5 m (shallow station) and -20 m (deep station) and at two different times of the day (12:00 and 18:00). Photosynthetic Active Radiation (PAR expressed in $\mu M/m^2/sec$) at the sampling time was 270 and 31 at -5 m (12:00 and 18:00, respectively), 86 and 5.39 at -20 m (12:00 and 18:00, respectively). Sampling was performed in the framework of a collaboration with other European and Australian research groups, on the multidisciplinary assessment of the *P. oceanica* meadows carbon budget (partially funded by the COST Action ES-0906; Mazzuca et al., 2013). Leaf tissue from two adult shoots collected at 5 m depth (12:00 and 18:00) and two at 20 m depth (12:00 and 18:00) was cleaned from epiphytes, stored in RNA later solution (Life Technologies) overnight at 4°C and then transferred at -20°C. Total RNA from *Posidonia* leaves was isolated using the Aurum Total RNA mini kit (Bio-Rad) following the manufacturer's instructions. The quality of total RNA was checked using the NanoDrop Spectrometer (ND-1000 Spectrophotometer, Peqlab) and the Agilent 2100 Bioanalyzer (RNA Nano Chip, Agilent).

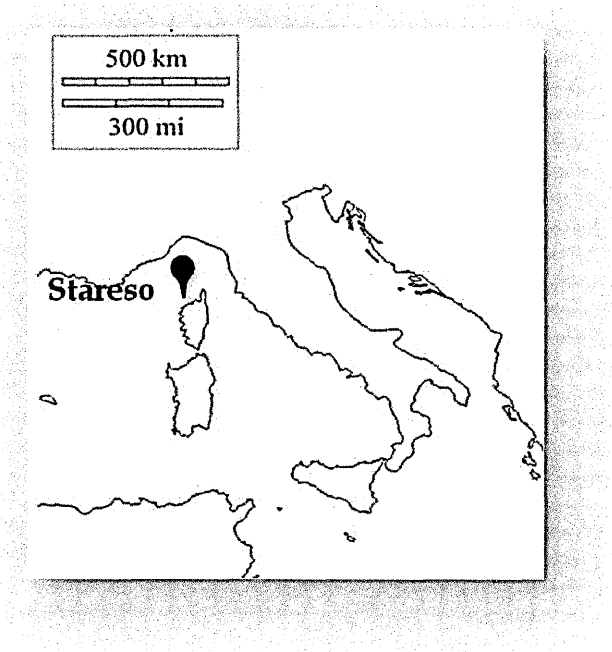


Figure 5.1. Geographical position of *Posidonia oceanica* sampling. Samples were collected at two different depths, -5 m (shallow station) and -20 m (deep station)

5.3.1-Transcriptome sequencing, assembly and annotation

High quality RNA samples were provided to the Genomic Research Centre of Fiorenzuola d'Arda, PC (Italian Agricultural Research Council - CRA) for library construction and sequencing. Four library were constructed starting from 3.5 µg of total RNA using the TruSeq RNA Sample Preparation Kit. Sequencing of the libraries was performed on the Illumina Genome Analyzer IIx instrument (Illumina, San Diego, California, USA) in two multiplexing runs. After de-multiplexing, high quality reads were used for the *de novo* assembly CLCbio Genomics Workbench (CLC Bio). Reads from each sample were then mapped on the assembled contigs and gene expression levels were estimated as FPKM (expected fragments per kilobase of transcript per million fragments

sequenced,Trapnell et al., 2010). Assembled transcriptome data were annotated using Blast2GO Pro software (Conesa et al., 2005). The KEGG annotation service KAAS was used for pathway annotation (Moriya et al., 2007). To classify the function of contigs, GO assignment was performed. Gene Ontology (GO) is a gene functional classification system which offers a dynamic-updated controlled vocabulary and a strictly defined concept to comprehensively describe properties of genes and their products in any organism (Ashburner et al., 2000).

5.3.3-Differential expression analysis

Expression values of GO classes were calculated by summing the relative counts for each gene in each sample associated to the specific GO class. The measures were plotted using R software (R Development Core Team, 2008). Outliers were not drawn.

GO enrichment analysis was performed with Blast2GO Pro (Conesa et al., 2005) comparing the 4 conditions: "day, shallow vs. deep", "night, shallow vs. deep", "shallow, day vs. night" and "deep, day vs. night". Generally, the best strategy to have accurate estimate of the overrepresented terms is to set up biological replicates for sequencing, but it greatly increases the experimental costs. In this study, to overcome the fact that only one biological sample for each condition was examined, a threshold FPKM ≥5 in selecting the lists of contigs to compare was applied. For each condition, the significance of over-representative GO terms was established with FDR ≤0.05.

5.4-Results

5.4.1-Illumina sequencing and de novo assembly

In order to characterize the transcriptome of *Posidonia oceanica* and generate a broad survey of genes associated with the response to different light and temperature conditions, RNA was isolated from plants living at different depth (shallow and deep) and at different times of the day (day and night). Four libraries were prepared and sequenced using the Illumina paired-end sequencing technology. Two multiplexing sequencings were ran, each containing two samples. After de-multiplexing, a total of 207,496,120 reads and 16,297,000 Mbp were generated. After data pre-processing and trimming, a total of 143,213,645 high quality reads were obtained and used for *de novo* assembly (Table 5.1).

Table 5.1. Sequence and assembly statistics.

Raw sequencing reads	
Number of reads	207,496,120
Total size (bp)	16,297,000
Pre-assembling	
Reads after trimming	143,213,645
Average read length (bp)	92,16
Assembly statistics	
Contigs	94,458
Number of contigs after filtering	53,001
Average lenght of contigs (bp)	1,071
Minimum length (bp)	140
Maximum length (bp)	24,532

According to the overlapping information of high-quality reads, a total of 94,458 contigs were obtained. After filtering for quality, 53,001 contigs were generated, whose lengths ranged from 140 bp to 24,532 bp, with an average length of 1,071 bp (Figure 5.2). To assess the accuracy of reconstructed contigs, reads for each sample were mapped back to the assembled reference transcriptome based on the pooled data: 27,906,075 reads, 25,820,509 reads, 26,792,477 reads, and 29,224,356 of the mapped reads were properly paired when aligning the "deep night", "shallow night" , "shallow day" and "deep day" samples, respectively, to the assembled reference transcriptome. The expression of contigs, quantified in FPKM, varied from 0 to 11,041,308 FPKM for deep night sample, from 0 to 10,766,149 for deep day sample, from 0 to 10,212,569 from shallow night and from 0 to 6,313,383 in shallow day.

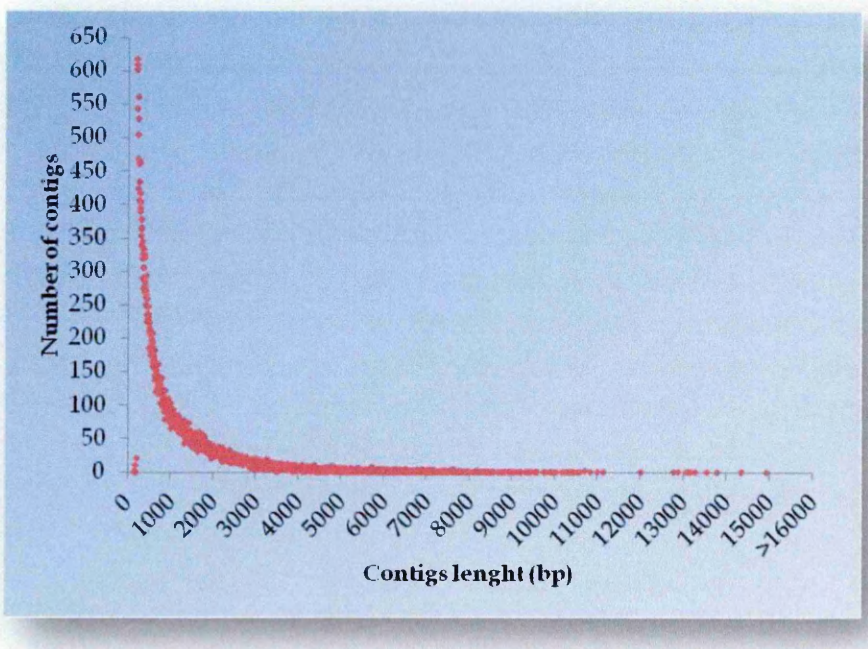


Figure 5.2. Distribution of contigs lenght (bp).

5.4.2-Sequence annotation

The annotation performed in Blast2GO Pro resulted in 18,485 (34.88%) annotated contigs, 1,667 contigs (3.14%) which provided a Blast result, 28,265 contigs without blast hits (53.32%), 93 (0.17%) contigs without blast results, and 4,501 (8.49%) contigs with mapping results.

Annotated contigs were assigned to one or more gene ontologies. Figure 5.3 shows the annotated contigs distribution regarding the number of GOs to which they were assigned. The number of GO terms per contigs ranged from 1 to 51, with the majority of the contigs mapped to 1 to 7 GO terms (Figure 5.3).

In total, 76,277 GO terms were retrived and classified according to their biological process, molecular function and cellular component. The distribution of annotated contigs, for each category, under different GO levels (Figure 5.4) reflects that they are concentrated in levels 2 to 6, indicating a good accuracy of annotation.

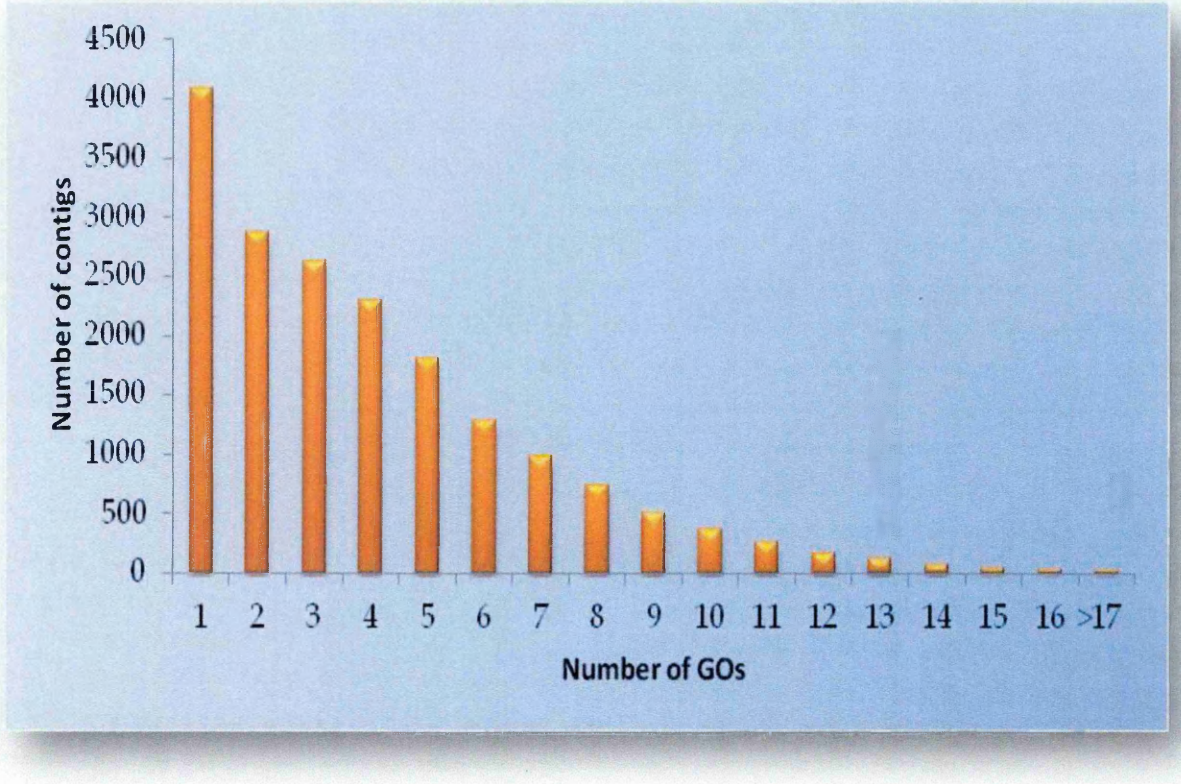


Figure 5.3. Number of GO terms distribution in the annotated contigs.

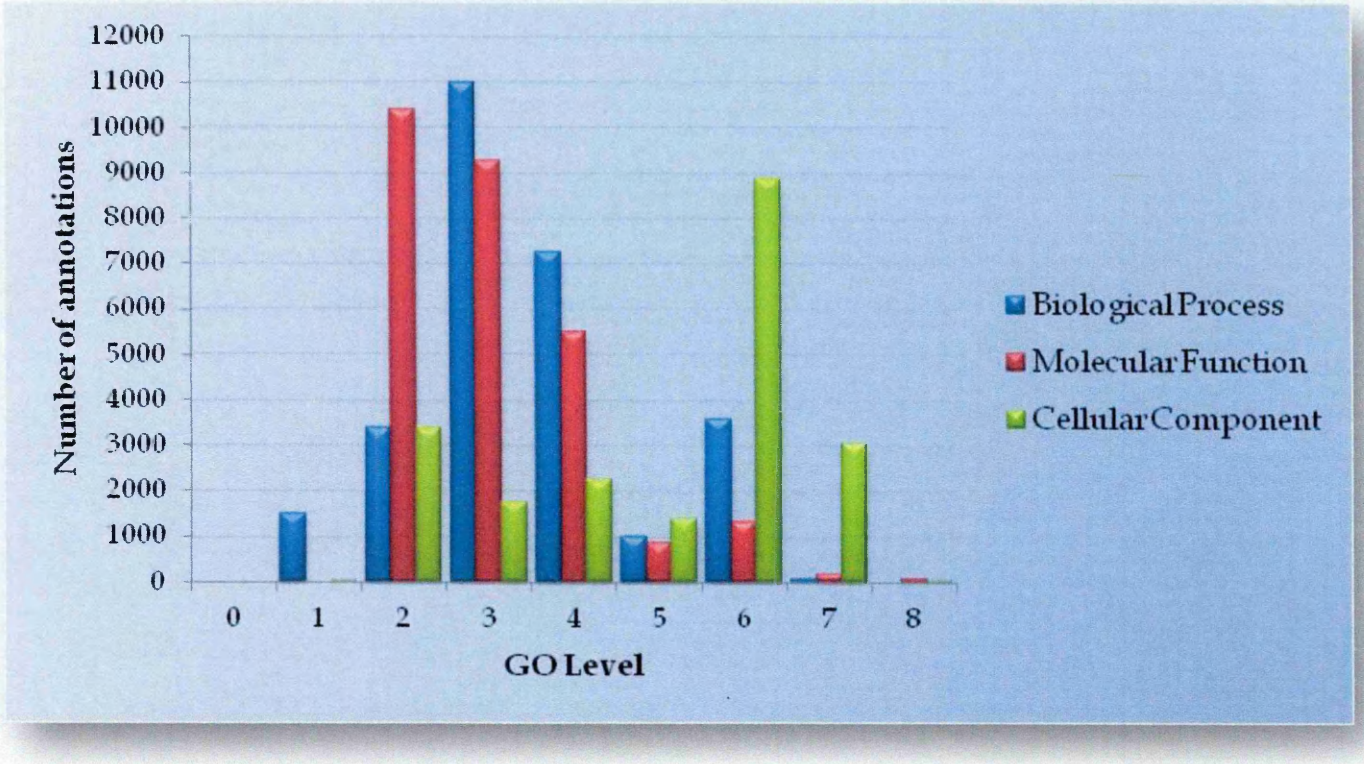


Figure 5.4. Distribution of annotated unigenes under different GO levels of each category.

Contigs were further classified into functional sub-categories within the three main categories (Figure 5.5, 5.6, and 5.7). Under the category of "biological process", metabolic (8,967, 35.23%) and cellular process (5,544, 21.78%) were more represented (Figure 5.5).

The sub-categories defined as *biological regulation* and *response to stimulus* were also well represented, with 2,254 (8.46%) and 2,138 (8.40%) contigs, respectively (Figure 5.5). Under the classification of "molecular function", binding (11,433, 50.06%) and catalytic activity (8,641, 37.83%) were the first and second largest sub-categories (Figure 5.6). The category antioxidant activity contained 75 contigs (0,33%). In the "cellular component" category, organelle (9,147, 35.88%) and cell (11,768, 46.17%) were the largest sub-categories, whereas only a few contigs were assigned to extracellular region and extracellular matrix (Figure 5.7).

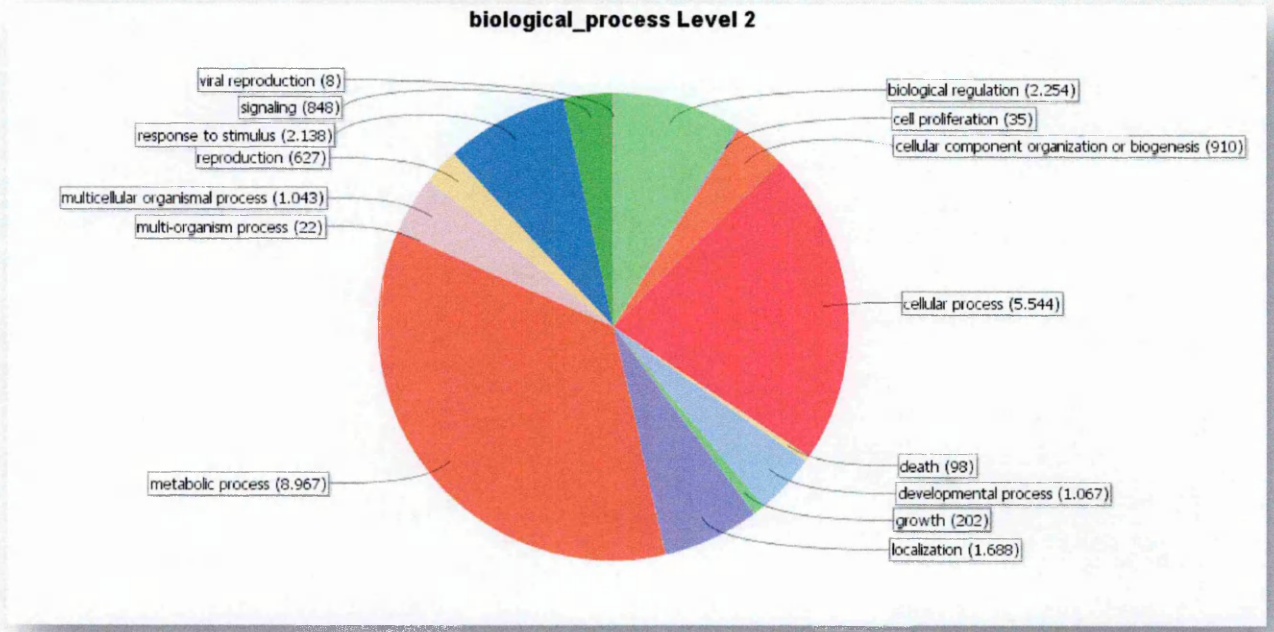


Figure 5.5. GO assignments (level 2) as predicted for their involvement in biological processes. The number of contigs assigned to each GO term is shown in brackets.

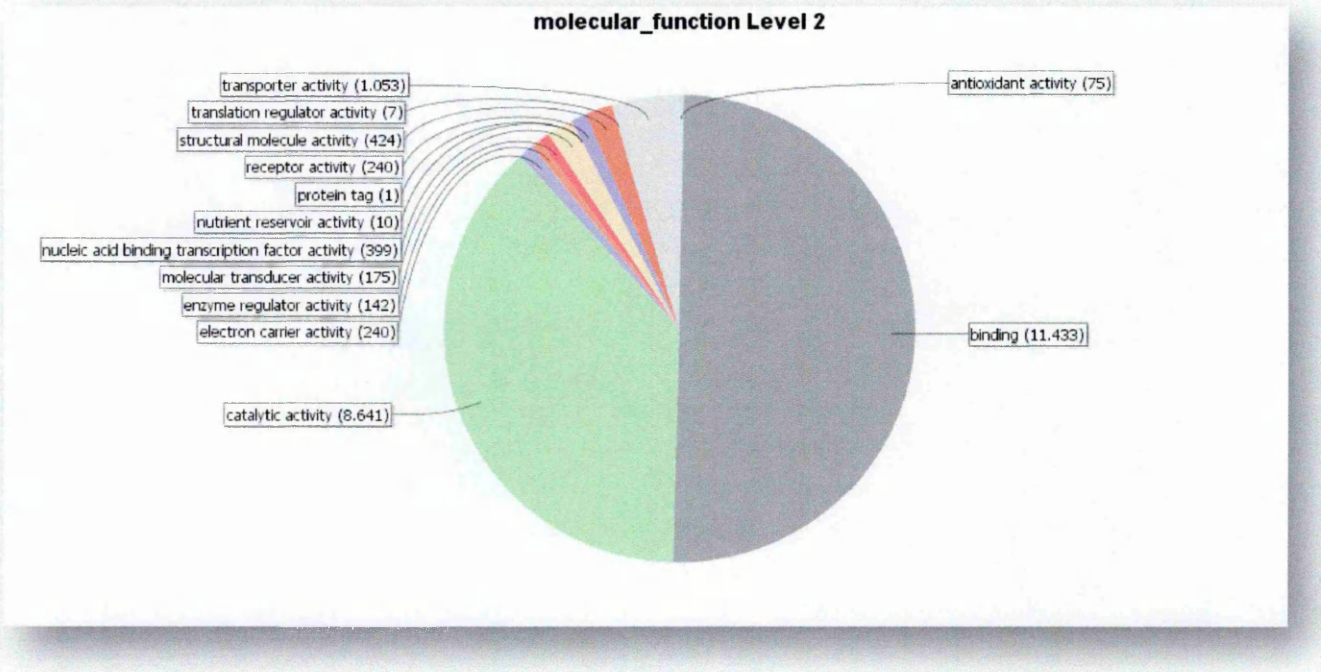


Figure 5.6. GO assignments (level 2) as predicted for their involvement in molecular function. The number of contigs assigned to each GO term is shown in brackets.

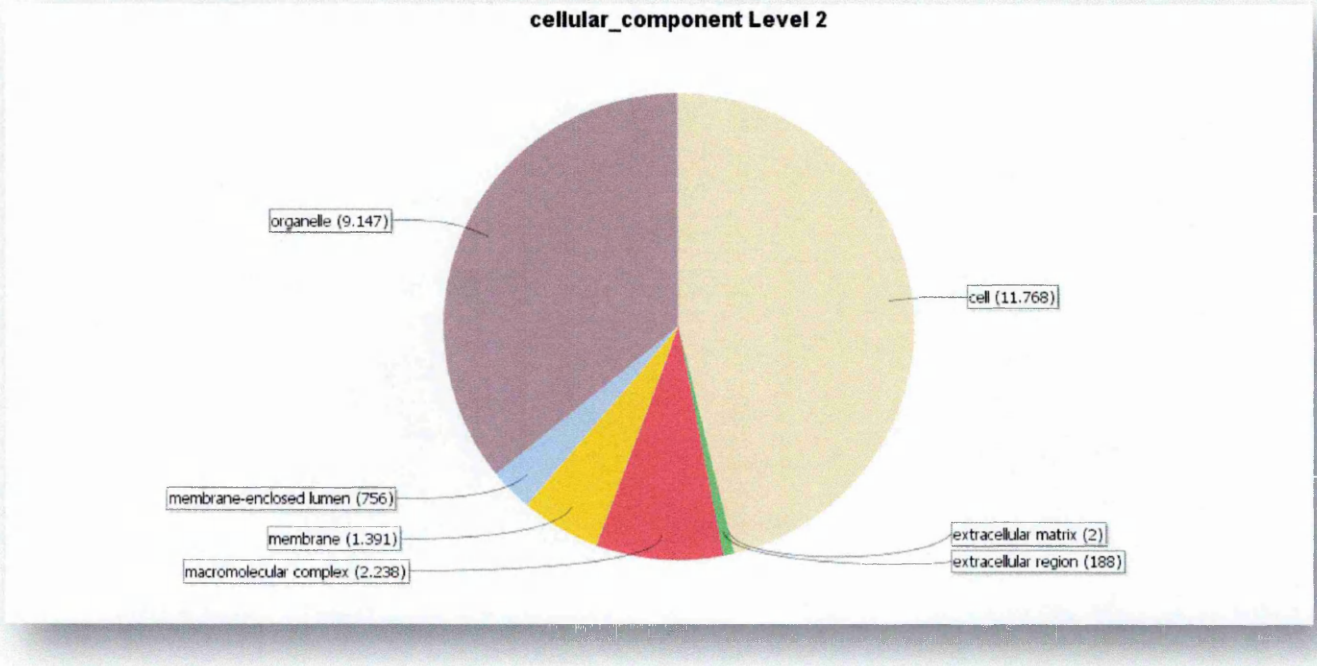


Figure 5.7. GO assignments (level 2) as predicted for their involvement in cellular component. The number of contigs assigned to each GO term is shown in brackets.

The KEGG Automatic Annotation Server (KAAS) identified 3,960 *Posidonia oceanica* contigs with KEGG orthologies (Moriya et al., 2007) belonging to 293 pathways. Among the others, 35 genes involved in the photosynthesis were identified. The Photosynthetic pathway as resulting from the KEEG server, is reported in Figure 5.8.

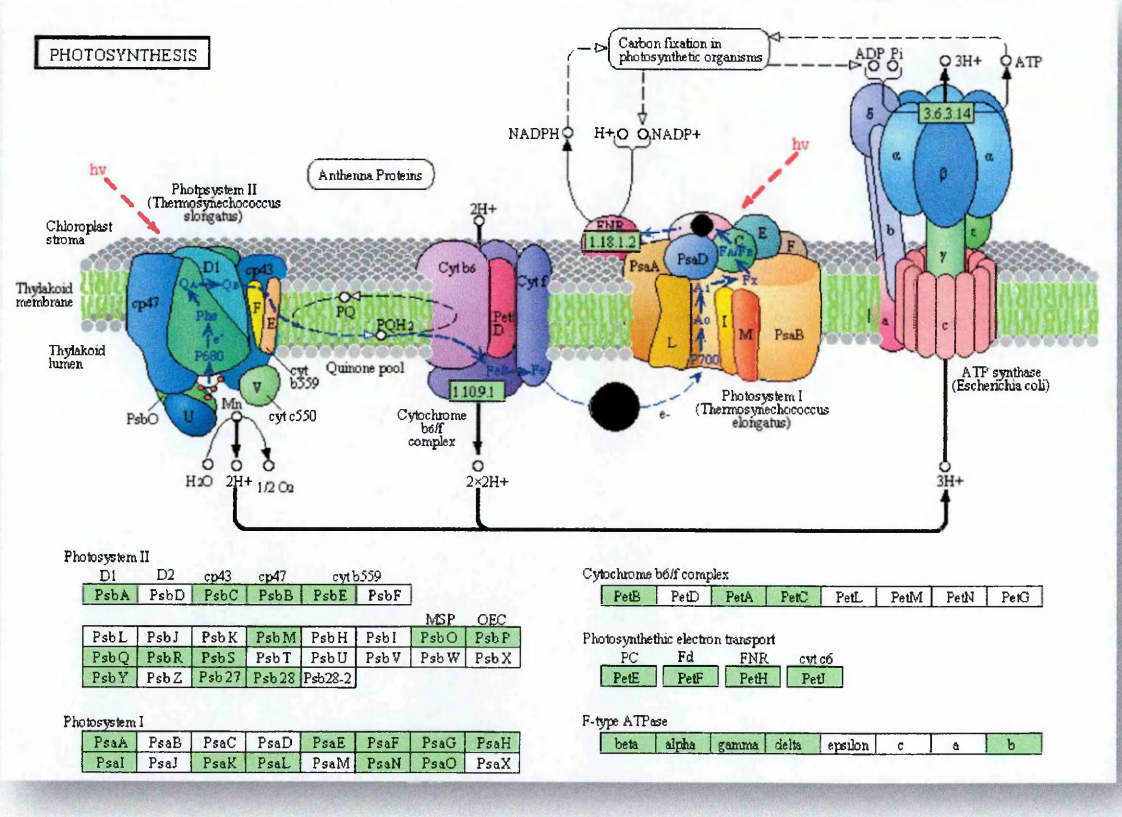


Figure 5.8. Reconstruction of Photosynthesis pathway in *Posidonia oceanica*. Genes present annotated in the *P. oceanica* transcriptome are highlighted in green.

5.4.3-Differential expression analysis of transcriptome data

Following GO classification, contigs gene expression levels were plotted for the four conditions. 146 groups of four boxplots were generated. Among the others, GO terms related to antioxidant activity highlighted the highest difference in gene expression in the "shallow vs. deep" condition as well as in the "day vs. night" condition (Figure 5.9) suggesting a stress imposed by high light and high temperature and the need for the plants to activate *ad hoc* defense systems.

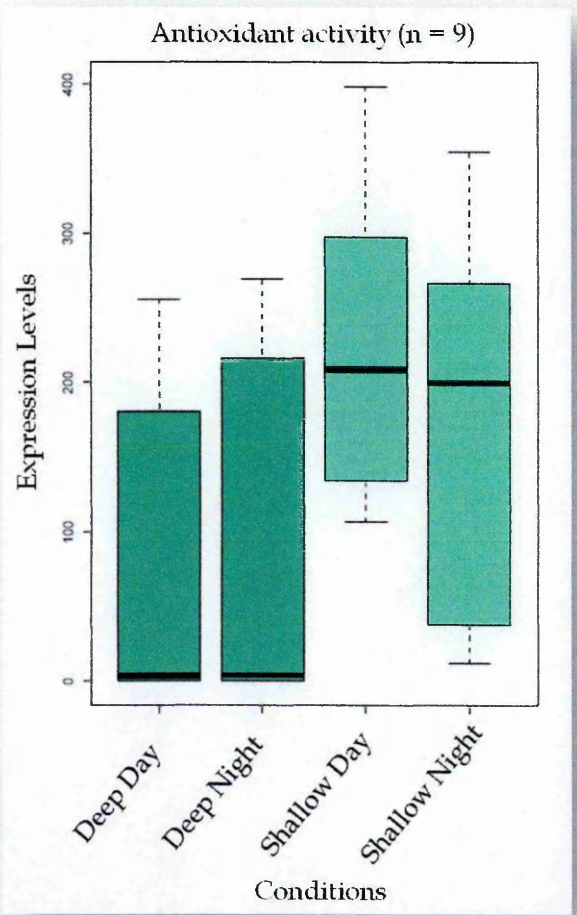


Figure 5.9. Distribution of expression levels related to antioxidant activity for the four conditions. n = number of contigs belonging to the GO category antioxidant activity.

To better highlight key biological processes involved in response to changes in light and temperature, a gene ontology (GO) enrichment analysis was performed on differentially expressed genes. The expressed transcripts were assigned to groups, following the GO categorization in biological processes, cellular components and molecular functions. Results are reported for each comparison performed.

Day - shallow vs. deep. In the "day shallow" condition, significantly enriched GO terms related to biological processes were included in the categories: "translation", "lipid metabolic process", "organelle organization", "cell cycle", "anatomical morphogenesis", "carbohydrate metabolic process" and "multicellular organismal development"(Table 5.2 and Figure 5.A1, in Appendix, for more details about all the enriched GO categories).

To better identify specific processes belonging to each category, a further investigation on representative genes of each category was performed, with the support of the KEGG pathways previously obtained (Table 5.2). For example, in "day shallow" condition an over-expression for transcripts related to defense and cell rescue mechanisms was observed. Within the lipid metabolism it was observed an over-expression for transcripts involved in the carotenoid biosynthesis, glicerolipid and glicerophospholipid metabolism and fatty acid biosynthesis. In the frame of the carbohydrate metabolic processes, the most representatives transcripts were related to the cellular respiration, to the sucrose and the starch metabolism and to the cell wall formation/remodeling. Also, over-

expressed transcripts were related to the regulation of cell cycle as well in the organization and functioning of the cytoskeleton.

In the "day deep" condition it was observed an enrichment for transcripts related to biological processes included in GO categories, such as "cell recognition", DNA metabolic process" and "cellular protein modification"(Table 5.2 and Figure 5.A1, in Appendix, for more details about all the enriched GO categories). In particular, (Table 5.2) it was observed an over-expression of transcripts for components of signal transduction, that belong to the GO categories "cell recognition" and "cellular protein modification" by means of phosphorilation and dephosphorilation, of protein, lipid and carbohydrate catabolism and of transcriptional regulation. Also in this condition, within the GO enriched categories an over-expression of transcripts for molecules with antioxidant activity was present.. Lastly, it was observed the over-expression of transcripts involved in the process of retrotransposition, that belong to the GO category "DNA metabolic process".

Table 5.2. Specific processes involved in the Day - Shallow vs. Deep comparison. F, P, C = GO terms molecular function, biological process and cellular component, respectively.

Day - shallow vs. deep		
	Day shallow	
Process		GO category
Defense and cell rescue meachanisms	• Antioxidant activity (e.g. peroxidase, respiratory burst or NADPH oxidases, cytochrome P450 CYP86B1, glutation-s-transferase) • DNA and Protein repair (e.g. heat shock protein 70, 40, DNA damage toleration protein DRT100) • Carotenoid biosynthesis (e.g. ABA2, xanthoxin dehydrogenase)	Antioxidant activity (F), Lipid metabolic process (P), Cell cycle (P), Extracellular region (C), Protein binding (F)
Gene expression	• Transcriptional regulation (e.g. nuclear transcription factor Y, HD-Zip transcription factors, MYB, EREBP, ANT, GATA, AP2ERF, BHLH-like, WRKI)	Protein complex (C), Protein binding (F), Multicellular organismal development (P)
Translation	• Ribosome assembly (e.g. ribosomal proteins)	Translation (P),

		Ribosome (C), Structural molecule activity (F), Cell cycle (P), Protein complex (C), Protein binding (F)
Hormone signaling	• Hormone signal transduction (e.g. auxin efflux carrier, protein brassinosteroid insensitive, ABA responsive element binding factor)	Cell cycle (P), Anatomical structure morphogenesis (P), Multicellular organismal development (P), Protein complex (C), Protein binding (F)
Lipid, protein and carbohydrate metabolism	• Glicerolipid and glicerophospholipid metabolism, (e.g. glycerol-3-phosphate acyltransferasebeta hexosamminidase, phospholipase A2-like) • Fatty acid biosynthesis (e.g. acycl acyl carrier protein thioesterase • Glycolysis/gluconeogranes, (e.g. piruvate kinases, aldose 1 epimerase,) • Carbon fixation (e.g.. rubisco, glycolate oxidase) • TCA cycle, (e.g. 2-oxoglutarate E1 component) • Starch and sucrose metabolism (e.g. sucrose synthase and cell wall invertase, starch branching enzyme, , trehalose phosphate synthase, beta glucosidase) • Cell wall formation/remodeling (e.g. callose synthase, cellulose synthase, protein COBRA, purple acid phosphatase, proteinases, extensions)	Lipid metabolic process (P), Carbohydrate metabolic process (P), Cell wall (C), Extracellular region (C), Protein complex (F),

		Protein binding (F)
Cell cycle and cytoskeleton	• Cell cycle regulation (e.g. calmodulin, cyclins, cyclin-dependent kinases, retinoblastoma related proteins, protein regulator of cytochinesis, BUB1, spindle assembly checkpoint factors, centromeric proteins, targeting protein for Xklp2, DNA replication initiation factor, condensins, nucleosome assembly factors, sister chromatid separation (Aurora), growth regulating factor 4, ovate family protein, maternal affect embryo arrest protein) • Cytoskeleton organization and functioning (e.g. tubulin, microtubule associated proteins, kinesin, myosins, dynein and dynamin)	Cell cycle (P), Cytoskeleton (C), Motor activity (F), Organelle organization (P), Structural molecule activity (F), Protein complex (C), Protein binding (F), Anatomical structure morphogenesis (P)
	Day deep	
Signal transduction	• Protein phosphorilation and dephosphorilation (e.g. Receptor like-kinases, component of MAP kinase signal transduction, phosphatases, , calcium/calmodulin dependent protein kinases)	Protein kinase activity (F), Cell recognition (P), Carbohydrate binding (F), Cellular protein modification (P)
Defense and cell rescue mechanisms	• Antioxidant activity (e.g. glutathione-s-transferase, cytochrome P450 monoxigenase CYP704G7, homogentisate phytyl transferase, homogentisic acid geranyl geranyl transferase, geranyl geranyl pyrophosphate synthase)	Plastid (C)

Gene expression	• Transcriptional regulation (e.g. zinc finger proteins WRKY, NAC and MYB BHBL, families and heat shock transcription factors)	DNA binding (F), Plastid (C)
Lipid, protein and carbohydrate metabolism	• Breakdown of protein, lipids and polysaccharides in chloroplasts (e.g.. carboxipeptidases and specifically for cystein protease, and a component of the ubiquitination system, beta amylase and lipoxigenases, pheophorbide a oxygenase, lethal leaf hot spot1-like protein, ubiquitination systems)	Plastid (C), Cellular protein modification (P)
DNA modification	• Retrotransposition (e.g. Ty1-copia and Ty3-gypsy group retrotransposons, gag, pol and int)	DNA metabolic process (P), DNA binding (F)

Night - shallow vs. deep. In the comparisons "night - shallow vs. deep", an over-representation of GO categories related to "ion transport", "lipid metabolic processes" and "carbohydrate metabolic processes" in the shallow condition was observed (Table 5.3 and Figure 5.A2, in Appendix, for more details about all the enriched GO categories). The main processes (Table 5.3) are related to the nutrient recycling and defense mechanisms. Within the lipid and carbohydrate metabolism, it was observed an enhanced hydrolytic activity and processes related to the starch and sucrose metabolism and amino sugars and nucleotide sugars metabolism, glycolysis and fatty acid biosynthesis.

In the night deep condition the over-representation of the GO categories related to biological processes "translation", "cell recognition", "cell cycle", "cytoskeleton organization" was present (Table 5.3 and Figure 5.A2, in Appendix, for more details about all the enriched GO categories). The specific processes in this condition are related to the signal transduction (that are included in the GO category: "cell recognition"), ribosome assembly and regulation of translation (that are included in the GO category: "translation"). Over-expressed transcripts are also involved in the regulation of cell cycle and in the organization and functioning of the cytoskeleton (that are included in the GO categories "cell cycle" and "cytoskeleton organization").

Table 5.3. Specific processes involved in the Night - Shallow vs. Deep comparison. F, P, C = GO terms molecular function, biological process and cellular component, respectively.

Night - shallow vs. deep		
Process	Night shallow	GO category
Defense meachanisms	• Antioxidant activity (e.g cytochromep450, peroxidases, respiratory burst oxidases)	Antioxidant activity (F), Cytoplasmic bounded-vescicle (C)
Protein, lipid and carbohydrate metabolism	• Breakdown of protein, lipids and polysaccharides, (e.gpectin methylesterase, lipase, phospholipase, galactosidase, polygalacturonase, proteinases) • Fatty acid biosynthesis (e.g. 3-ketoacyl-CoA synthase, fatty acyl-ACP thioesterase B) • Starch and sucrose metabolism and amino sugars and nucleotide sugars metabolism (e.g. sucrose synthase, cell wall invertase), • Glycolysis (e.g. pyruvate decarboxylase, fosfofructokinase)	Enzyme regulatory activity (F), Hydrolase activity (F), Cell wall (C), Carbohydrate metabolic process (P), Lipid metabolic process (P)
Nutrient recycling	• Transport (e.g. magnesium, sodium, potassium, nitrate transporters, and ABC transporters, cytoplasmic membrane bounded vescicle, lipid transfer protein)	Ion transport (P), Lipid binding (F), Cytoplasmic membrane

		bounded-vesicle (C), Hydrolase activity (F)
	Night Deep	
Signal transduction	• Triggering and Phosphorilation (e.g. calmodulin, receptor-like kinase, protein kinases)	Protein kinalse activity (F), Cell recognition (P), Carbohydrate binding (F)
translation	• Ribosome assembly (e.g. ribosomal proteins) • Translation regulation (e.g. translation initiation and elongation factors)	Translation (P), Ribosome (C), Structural molecule activity (F), Translation factor activity-nucleic acid binding (F), Cell cycle (P), Cytosol (C)
Cell cycle and cytoskeleton	• Cell cycle regulation (e.g. calmodulin, mitotic spindle chechkpoint components and nucleoside diphosphate kinase, ribosomal proteins) • Cytoskeleton structural component (e.g. actin, tubulin, profilin)	Cell cycle (C), Cytoskeleton organization (P), Cytosol (C)

Shallow - day vs. night. At shallow depth, it was observed an over-representation of the following GO categories related to biological processes: "cellular component organization", translation", "lipid metabolic process", "response to stimulus", "regulation of biological process" and "catabolic process" in the day condition (Table 5.4 and Figure 5.A3, in Appendix, for more details about all the enriched GO categories). The analysis of the specific transcripts belonging to these categories (Table 5.4) showed an over-expression of transcripts involved in the translational and transcriptional regulation (that explain the presence of the GO category: "regulation of biological processes", as one of the over-expressed), and signal transduction, in cell defense mechanism and hormone signaling (that explain the presence of the GO category: "response to stimulus", as one of the over-expressed). Within the frame of "catabolic process", an over-expression of transcripts was observed for proteins involved in the respiration process. The over-expression of transcripts involved in the regulation of cell cycle and in the organization and functioning of the cytoskeleton explain the over-representation of the GO category "cellular component organization".

In the "shallow night" condition, we have an over-representation of the GO categories "DNA metabolic process", "nucleic acid binding" and "protein kinase activity" (Table 5.4 and Figure 5.A3, in Appendix, for more details). The over-expressed transcripts which belong to these categories are involved in signal transduction (GO category: "protein kinase activity"), transcriptional regulators and retrotransposition (GO categories: "nucleic acid binding" and "DNA metabolic process"; Table 5.4).

Table 5.4. Specific processes involved in the Shallow - Day vs. Night comparison. F, P, C = GO terms molecular function, biological process and cellular component, respectively.

Shallow Day vs. Night		
Process	Shallow Day	GO category
Signal transduction	• Triggering and Phosphorilation (e.g. calmodulin, protein kinases)	Response to stimulus (P), Cell periphery (C), Extracellular region (C)
Defense and cell rescue mechanisms	• Oxidant activity, (e.g. glutathione-s-transferase, peroxidases and respiratory burst oxidase, alternative oxidase, cytochrome p450) • DNA and protein repair (e.g. heat shock protein 70-90-83-like, DNA-3-methyladenine glycosilase i, chromatin assembly factor-1,DNA damage toleration protein DRT100) • Redox state regulation (e.g. proton gradient regulation 5, PGR5, and ferric reductae-like transmembrane component also glutaredoxin-c9) • Carotenoid biosynthesis (e.g. ascisic acid 8 hydrolase, beta-carotene hydroxilases) • Flavonoid biosynthesis (e.g. fructokinase 3)	Response to stimulus (P), Cell periphery (C), Lipid metabolic process (P), Nucleus(C)
Gene expression	• Transcriptional regulation (e.g. zinc finger proteins , including WRKY family)	Regulation of biological process (P),

		Nucleus (C)
Translation	• Ribosome assembly (e.g. ribosomal proteins) • Traslation regulation (e.g. translation initiation factors, elongation factors)	Translation (P), Structural molecule activity (F), Regulation of biological process (P)
Lipid, protein and polysaccharide metabolic process	• Breakdown of protein, lipids and polysaccharides (e.g. subtilisin-like serine protease, proteinase and ubiquitin conjugating enzymes, lipoxigenases, fatty acid hydrolases) • Glycolyis (e.g. glyceraldehydes-3-phosphate dehydrogenase, aldose-1-epimerase, phosphoglucomutase, piruvate dehydrogenase piruvate decarboxilase etc) • Kreb cycle (e.g. ATP-citrate synthase, isocitrate dehydrogenase). • Oxidative phosphorilation (e.g. F-type H+-transporting ATPase subunit beta)	Catabolic process (P), Lipid metabolic process (P)
Cell cycle and Cytoskeleton	• Cell cycle regulation (e.g. calmodulin, cyclin dependent kinases, mitotic spindle checkpoint proteins,DNA polymerase, chromatin assembly factors, histones, condensins, microtubule associated proteins, nucleoside diphosphate kinase) • Cytoskeleton organization and functioning (e.g. tubulin,profiling, myosin)	Structural molecule activity (F), Cellular component organization (P), Extracellular region (C), Cell periphery (C), Cytoskeleton (C), Nucleus(C)

Hormone signaling	• Hormone signal transduction (e.g. auxin response factor, two-component response regulator ARR-A family,abscisic acid receptor PYR/PYL family, SAUR family protein)	Response to stimulus (P), Nucleus (C)
	Shallow night	
Signal transduction	• Phosphorilation (e.g. receptor like kinases, component of MAP kinase signal transduction)	Protein kinase activity (F)
Gene expression	• Transcriptional regulation (e.g. MYB family, zinc finger protein)	Nucleic acid binding (F)
DNA process	• Retrotransposition (e.g. Retrotransposons, T3-GYPSY subclass, Ty1-copia subclass, integrase core domain containing protein, polyproteins, DNA topoisomerase, DNA polymerase)	DNA metabolic process (P), Nucleic acid binding (F)

Deep - day vs. night. In the day condition, it was observed an over-representation of GO categories related to biological processes such as "cellular protein modification", "lipid metabolism" and "carbohydrate metabolism"(Table 5.5 and Figure 5.A4, in Appendix for more details about all the enriched GO categories). In particular, it was observed an over-expression of transcripts involved in the signal transduction (GO category cellular protein modification), cell defense mechanism and nutrient recycling (Table 5.5). Within the lipid and carbohydrate metabolism, over-expressed transcripts were related to cell wall remodeling, starch and sucrose metabolism, fatty acid biosynthesis, and flavonoid biosynthesis (Table 5.5).

In the "deep night" condition, plants are enriched for transcripts involved in the process of "translation", "cytoskeleton organization", "anatomical structure development" and "cell cycle"(Table 5.5 and Figure 5.A4, in Appendix for more details about all the enriched GO categories). In this condition, within the GO category "translation", plants over-expresses transcripts involved in the ribosome assembly and in the regulation of translation, while within the GO categories "anatomical structure organization" and "cytoskeleton organization", plants over-expresses transcripts involved in the regulation of the cell cycle and in the organization and functioning of the cytoskeleton (Table 5.5).

Table 5.5. Specific processes involved in the Deep Day vs. Night comparison. F, P, C = GO terms molecular function, biological process and cellular component, respectively.

Deep Day vs. Night		
Process	**Deep Day**	GO category
Signal transduction	• Protein phosphorilation and dephosphorilation (e.g. protein kinases, phosphatases)	Protein kinase activity (F), Nucleotide binding (F), Cellular protein modification (P), Hydrolase activity (F), Cell wall (C)
Defense and cell rescue mechanisms	• Antioxidant activity (e.g. respiratory burst oxidases and cytocrome P450) • Protein, DNA repair (e.g. heat shock proteins, DNA damage toleration proteins, DNA methyadenine glycosylase) • Flavonoid biosynthesis (e.g. dyhydroflavonol-4-reductase)	Nucleotide binding (F), Lipid metabolic process (P)
Lipid and carbohydrate metabolism	• Cell wall remodeling (e.g. xiloglucan endotransglucosylase hydrolase, beta-glucanase, polygalacturonase, pectin esterase, pectin methilesterase, protease, proteinase) • Starch and sugar metabolism (e.g. beta-amylase, starch synthase, trehalose-phosphate phosphate like, UDP-glucose 4,6-dehydratase, UDPglucose 6-dehydrogenase) • Pentose phosphate pathway and Glycolysis (e.g. glucose -6-phosphate	Nucleotide binding (F), Endoplasmic reticulum (C), Lipid metabolic process (P), Carbohydrate metabolic process (P),

	dehydrogenase, piruvate decarboxilase) • Fatty acid biosynthesis (e.g. 3-ketoacyl synthase, fiddlehead like protein)	Hydrolase activity (F), Cell wall (C)
Nutrient recycling	• Nutrient transport (e.g. ABC transporters, nitrogen and potassium, zinc and iron transporters, sugar porter)	Transporting activity (F)
	Deep Night	
Translation	• Ribosome biogenesis and assembly (e.g. ribosomal proteins, DNA directed RNA polymerase I subunit, ribonucleoprtoein complex subunit) • Regulation of translation (e.g. translation initiation factor and translation elongation factor)	Nucleolus (C), Translation (P), Ribosome (C), Structural molecule activity (F), Translation factor activity, nucleic acid binding (F), Cytosol (C), Anatomical structure development (P)
Cell cycle and cytoskeleton	• Cell cycle regulation, (e.g. calmodulin, ribosomal proteins, DNA topoisomerase, splindle formation protein, microtubule associated proteins, nucleoside diphosphate kinase) • Cytosckeleton organization (e.g. actin, tubulin, formin-like protein, profilin, kinesin related protein)	Cell cycle (P), Anatomical structure development (P), Cytoskeleton organization (P), Cytoskeleton (C), Cytoskeleton protein binding (F)

5.5-Discussion

Transcriptome sequencing can provide abundant sequence information, as well shedding some light on the basic biological processes occurring in the cells (Birol et al., 2009; Wu et al., 2012). In this study, RNAseq technology was used to sequence, the *P. oceanica* transcriptome and to discover a large number of novel genes.

From Illumina sequencing 143,213,645 clean reads were produced. Because of the absence of reference sequences for *P. oceanica*, the *de novo* assembly was apply to connect these short reads to form longer expressed sequences, that resulted in the generation of 53,001 contigs. Of these, 18,485 contigs were successfully annotated (about 34.88% of the assembled contigs), suggesting their relatively conserved functions. Functional classification showed that this contigs covered every basic biological processes, and 3,960 of these contigs were also mapped into 293 KEGG pathways. Among the others, KEGG prediction results suggest that the photosynthesis related-genes were rather well conserved in *Posidonia.*

Currently, in the public seagrass database, Dr. Zompo, only 2,069 ESTs are available for *P. oceanica*. Thus, the ESTs generated in the present study represent a significant addition to the existing genomics resources of this specie.

RNA-Seq results also provided a measure of expression for each of the genes allowing for differential gene expression analyses. In this study, the response to different light and temperature conditions was assessed comparing the GO category enriched in four conditions: "day, shallow vs. deep", "night, shallow vs. deep", "shallow, day vs. night" and "deep, day vs. night". Following, the four comparisons performed, are discussed separately.

5.5.1-Day - shallow vs. deep

Day shallow. Shallow and deep plants, inevitably, experience different conditions of light and temperature, with shallow plants more exposed to heat and high light intensity. From this study, there are evidences that shallow plants are responding to this with the enhancement of their **antioxidant activity** by producing and scavenging ROS (Reactive Oxygen Species).

ROS are normally produced in basic cellular processes such as photosynthesis, respiration and photorespiration, but they increase in response to a variety of abiotic stimuli, including extreme temperature and high light. The imbalance in the generation and in the metabolism of ROS, with more ROS being produced than what are metabolized, can result in oxidative stress (Suzuki et al., 2011) that can cause severe damage to protein structures, can inhibit the activity of multiple enzymes of important metabolic pathways, and can result in oxidation of macromolecules including lipids and DNA. All these adverse events compromise cellular integrity and may lead to cell death (Fraire-Velázquez and Balderas-Hernández, 2013).

Most recent studies have focused on ROS as important signal transduction molecules (Figure 5.10). The key of using ROS as signaling molecules appears to be the capacity of cells to detoxify or scavenge them using a network of ROS scavenging enzymes. This network enables cells to maintain a nontoxic steady-state level of ROS, while allowing for the transient accumulation of ROS in particular subcellular locations, that act as signals (Suzuki et al., 2011).

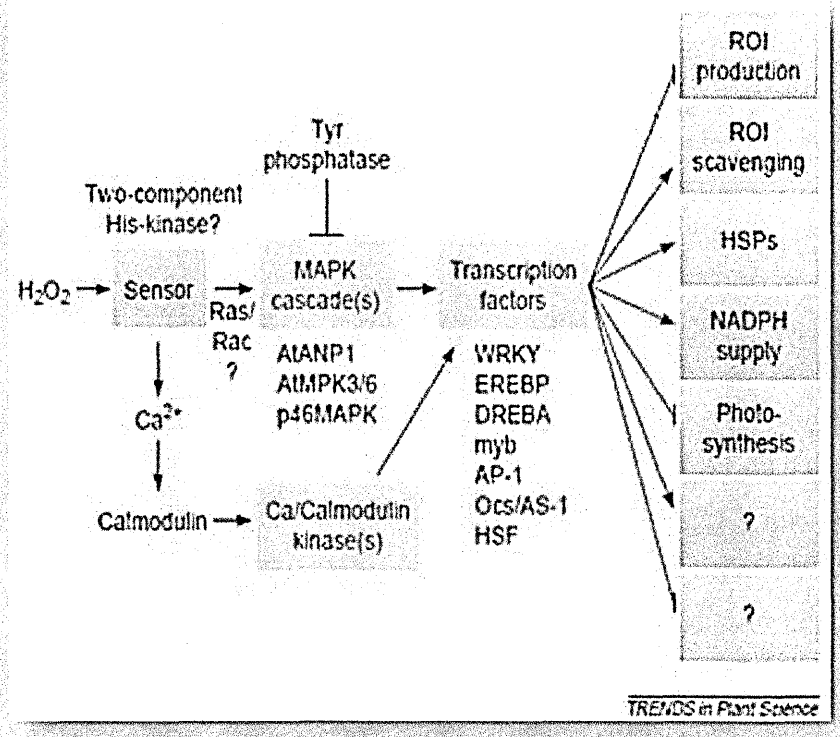


Figure 5.10. A suggested model for the activation of signal transduction events during oxidative stress (from Mittler, 2002). MAPK= mitogenactivated-protein kinase; ROI= reactive oxygen intermediate, HSPs = heat-shock proteins.

The elevation of cytosolic calcium concentrations has been shown to occur during oxidative stress (Mittler, 2002) and could explain the increase in the transcripts for calmodulin that was observed in the data. The increase in calcium, activate calcium dependent protein kinases and phosphatases as early step, with some enzymes potentially mediating downstream components such as other protein kinases/phosphatases and other effector proteins. A protein phosphorilation cascade that has been shown to be activated by H_2O_2 is a mitogen activated protein kinase (MAPK) cascade, that activates transcription factors, thereby modulation gene expression (Figure 10, Mittler, 2002).

Shallow plants showed an over-expression of transcripts for all the components of this cascade (refer to Table 5.2). Our data suggested that plants are activating defense mechanism against oxidative stress and **repairing mechanism for proteins and for DNA**. In addition, membranes, particularly plasma and chloroplast membranes, are sensitive to environmental stimuli, thus, maintaining the integrity and fluidity of membranes is of fundamental importance for plants to survive, for example, to high temperature stress. The over-expression of transcripts involved in the **metabolism of membrane components** (glicerolipid and glicerophospholipid metabolism and fatty acid biosynthesis) could suggest that plants are preparing the machinery to adjust the composition or the degree of unsaturation of the lipids of their membranes to maintain the integrity and the optimal fluidity, in response to change in temperature.

The increase in **translation** could explain the need for plants to accumulate defense-related proteins. Translation is a highly energy-demanding process and, consequently, it is one of the main targets to be inhibited in response to most types of cellular stresses. However, under conditions where global protein synthesis is severally impaired, some proteins mainly involved in homeostasis maintenance remain being synthesized as part of the mechanisms of cell survival. In plants, several studies have demonstrated that the modulation of mRNA translation is a key point in the adaptation of plants to different challenging conditions (Floris et al., 2009). Protein synthesis is a key step of gene expression and it is especially regulated at the initiation phase (Sonenberg and Hinnebusch, 2009), as results also for our data in which an over-expression of transcripts for translation initiation

factors was found. Here, the hypothesis is that plants are using translation mainly for proteins involved in cellular defense and rescue mechanisms.

In the day, the plants also showed an enhancement for **sucrose and starch metabolism**. By generating hexoses and their derivate, invertase or sucrose synthase-mediated sucrose metabolism provide (1) energy source to power cellular processes; (2) starting molecules convertible to numerous metabolites and building blocks for synthesizing essential polymers including starch, cellulose, callose, and proteins; and (3) a mechanism to reduce sucrose concentration at the unloading sites to facilitate its source-to-sink translocation, thereby preventing feedback inhibition on photosynthesis and sustaining carbon flow at the whole-plant level (Ruan, 2012). In addition, plants are also over-expressing transcripts for protein involved in the **cellular respiration** (glycolysis/gluconeogranes, TCA). Respiration breaks down the products of photosynthesis to liberate energy (ATP and reducing equivalents) and carbon skeletons necessary for growth, cellular maintenance and active transport (including ion uptake from soils; Atkin and Tjoelker, 2003). The over-expression of Rubisco small subunit and glycolate oxidase, two component of the photorespiration process, could suggest the photoprotective role of this process, to dissipate excess energy at high irradiance levels (Kozaki and Takeka, 1996). In addition there is evidence that plants are over-expressing transcripts for the **regulation the cell cycle** progression as well as for the **organization and functioning of their cytoskeleton**. In addition, the construction of a new cell naturally requires the construction of a new cell wall, with the deposition of different cell wall components, that could explain the over-expression of transcripts for component of the cell wall formation proteins.

Data also provide the evidence that plants may be going to expand their cells as resulted from an enrichment of proteins involved in the remodelling of wall polimers and in transcripts for proteins involved in cross-linking the wall, as revealed by the over-expression of purple acid phosphatase, implicated in the cell wall remodelling by expansion. In addition there is evidence that cells are elongation. The overrepresentation of protein COBRA also support this hypothesis, this protein is involved in the regulation of cell shape during elongation (Schindelman et al., 2001; Borner et al, 2002). Also in this case, overexpressed component of the cytoskeleton may be required for the structure of expanding cells. The presence of an auxin afflux carrier, could suggest a role of the auxin in the activation of cell cycle. Auxin has been implicated in cell division as well as cell elongation. Evidence from many studies including those in *Arabidopsis* suggest that certain times of the day are important for plant growth. Temperature signals may act to modify auxin synthesis or distribution in the plant, and this could represent a growth-regulation mechanism (Penfield, 2008). Also, Gray et al (1998) showed that high temperature promote an increase in auxin levels and consequently auxin transport resulting in increased hypocotyls elongation. Here, higher temperature at shallow water could be responsible for the acceleration and encourage the development and growing, modifying the plants architecture.

An explanation for the cell expansion and elongation is possibly related to the tropism. Plants use light not only as an energy source for photosynthesis but also as an environmental signal and respond to its intensity, wavelength, and direction. Light affects the direction of growth by controlling the distribution of auxin in the shoot tip, so that more auxin gathers on the side furthest from the light. Therefore

more auxin passes down the shaded side of the shoot where it increase the rate of cell elongation and cause a growth curvature towards the light (Beckett, 1986). Auxin for example functions for increasing the plasticity of plant cell wall and for enlargement and shaping, influences the expression of specific genes involved in the growth and has a role in the stimulation of cell cycle. The overrepresentation of transcript for auxin efflux carrier in our data could be relate to the necessity for the plants to distribute this hormone in the leaf. On the other hand, auxin could be the signaling for the cell cycle.

There is also another explanation about the role of the cytoskeleton that, in addition to the relation with the cell cycle, could be related to the high light and to the orientation and motility of the chloroplast. In fact, plant cells respond to a variety of internal and external stimuli with rapid and dramatic rearrangements of their cytoplasm, that are often mediated by a dynamic cytoskeleton. There is a growing evidence that signal transduction cascades converge on these cytoskeleton proteins during many fundamental cellular responses. The orientation and motility of chloroplasts are precisely regulated in response to the angle, wavelength, and fluence rate of the light. Plant cells optimize the use of light for photosynthesis, by maintaining control over plastid position. Conversely, when fluence rates are too high, plastids can be positioned to avoid or minimize photodamage. Following its detection by photoreceptors in the cytoplasm, the light signal is rapidly translated into changes in the behavior and localization of the actin cytoskeleton, through as yet poorly defined signal transduction events (Staiger, 2000). In addition, the stimulation of elongation growth is also the most dramatic shade avoidance response, that is a strategy that plants evolved in

response to competition for light in respect to plants growing in close proximity (Morelli and Ruberti, 2002). This could be true for *Posidonia oceanica* that forms dense meadows especially at shallow sites, where plants grow near each other.

Day Deep. Also the deep plants (refer to Table 5.2) are involved in the **signal sensing** and **signal transduction,** with the activation of transcription factors for genes involved in the **cell defense and rescue from the oxidative stress,** such as heat shock proteins, cytochrome P450, radical scavenger metabolites (i.e glutathione-s-transferase) and components of the tochopherol biosynthesis. The over-expression of transcripts for proteins involved in the **catabolism** of chlorophyll and macromolecules in the chloroplasts, such as protein and membrane lipids, may be a response of the plants to prevent the catastrophic damages resulting from an excess of reactive oxygen species.

Another aspect that characterizes the response of the plants to live in deep waters is an increase in transcripts for **retrotransposons**. The activity of the retrotransposons is tightly controlled (Vukich et al., 2009) because, for example, the insertion of transposable elements (TEs) can modify the expression or the coding capacity of genes and can be extremely deleterious (Casacuberta et al., 2003). The regulation of transposons transcription proceeds with the binding of transcription factors to retrotransposons promoter. In our data, we found an overrepresentation of Myb and WRI transcription factors that could have a role in the activation of retrotrasnposons (Kumar and Bennetzen, 2000). While the transcription of most of the active plant elements characterized to date is largely quiescent during normal development, a growing body of evidence shows that

abiotic and biotic stresses are major factors in their transcriptional and transpositional activation (Grandbastien, 1998; Kumar and Bennetzen, 2000; Ito et al., 2013) The variability that the movement of TEs generates could also help the genome to rapidly evolve when exposed to conditions thatit is not well adapted, as initially proposed by McClintock (1984). Aside from their mobility, there are further ways in which TEs can modify gene expression and play a role in stress adaptation. For example, TEs generate a vast repertoire of small RNAs that can regulate subsets of stress-related genes or act as general transcriptional repressors upon stress. These small RNAs, being mobile, could influence gene expression in distant targets, which would be especially important in the stress adaptations of plants (Mirouze and Paszkowski, 2011). It is possible that in the deep plants, retrotrasposons provide a response and a defense to the temperature fluctuations, far from the optimal values for plant growth.

5.5.2-Night - shallow vs. deep:

Nigh shallow. In the "night shallow" condition (refer to Table 5.3), plants are involved in the **breakdown of molecules** (i.e. sucrose) accumulated during the day with photosynthesis and in their transport from the source to the sink tissues. The over-expression of transcripts for proteins involved in the **degradation of starch** and in the **cleavage of sucrose** could be the way for the plants to provide substrates for glycolsis/respiration. **Transport of nutrients or molecules** is also enhanced (magnesium, sodium, potassium, nitrate and ABC transporters lipids and sugars), as well as transcripts for cytoplasmic membrane bounded vesicles to

store waste products or transport substances. The over-expression for transcripts encoding for proteins involved in the cell wall could explain a cell remodeling to allow the transport of substances between cells.

Night Deep. In "night deep" condition (refer to Table 5.3), plants show evidence that are growing, preparing the **cytoskeleton organization** and increasing their **translation machinery** to synthesize proteins involved in the **cell cycle**.

5.5.3-Shallow - day vs. night

Shallow day. In the day, shallow plants (refer to Table 5.4) are involved in the response to high light and high temperature and a similar response to that showed in the comparison "day - shallow vs. deep" was observed. As discussed above, plants are responding to the **oxidative stress** due to light and temperature stress by activating a **signal transduction** cascade involving the ROS signaling and thus, the activation of the MAPK cascades that regulates the gene expression by activating stress response genes that protect plants from the environmental cues (Desikan et al., 1998, 1999, 2000; Kovtun et al., 2000). Also the observed increase in transcript for WRKY family **transcription factors**, has been showed to be induced from oxidative stress (Chen et al., 2012). Moreover, there is evidence for the activation of mechanisms for detoxification and protection, as showed by the over-expression of transcripts for antioxidant enzymes (glutathione-s-transferase, peroxidases and respiratory burst oxidase and alternative oxidase), for protective genes involved in **protein repair** (heat shock proteins, 70-90-83-like,) in the **redox homeostasis** and in the **DNA repair** mechanisms (DNA-3-methyladenine

glycosilase i, chromatin assembly factor-1,DNA damage toleration protein DRT100). Plants are responding also to minimize damages to the photosynthetic apparatus by up-regulating the expression of transcripts involved in the regulation of the cyclic electron flow around PSI, a mechanism essential for photoprotection (i.e. proton gradient regulation 5, PGR5, and ferric reductase-like transmembrane component and glutaredoxin-c9). Within the lipid metabolism, we have components of the **biosynthesis of carotenoids** (i.e. ABA2, abscissic acid 8 hydrolase), that protects chlorophyll from photodamage, and of the **biosynthesis of flavonoids** (i.e fructokinase 3), that also have an antioxidant role. There is evidence for a **catabolic process** whereby reserve food materials like carbohydrates, fats and proteins are subjected to hydrolysis by specific enzymes. The enhanced hydrolytic activity provides the components for respiration processes. Over-expression for component of **glycolyis** and for component of the **Kreb cycle** and **oxidative phosphorilation** were observed. The over-representation of transcript related to glycolysis could be due to fact that NADPH produced by the glycolysis is necessary for the biosynthesis of fatty acid related to the membrane lipid generation and antioxidant systems such as acerbate-gluthatione cycle. In addition, also the oxidative phosphorilation (together with glycolisis) could provide ATP to meet the needs for fatty acid synthesis.

Shallow night. In the night, plants are mainly involved in the DNA metabolic process represented by transcripts for **retrotransposons** (refer to table 5.4). As said before, LTR retrotransposons seems to play an important role in genome reorganization induced by environmental challenges. The environmental stresses

cause an epigenetic activation of mobile elements dispersed throughout the genome (Mansour, 2007).

5.5.4-Deep - day vs. night

Deep day. Also the deep plants (refer to Table 5.5) during the day are activating ROS signaling cascades, as evident from the over-representation for transcripts encoding molecules involved in the **ROS signaling** and in the **repair of protein, DNA and lipids**. The fine regulation of this cascade is showed by the enrichment of transcripts for **phosphorilation and dephosphorilation**. There is evidence for a **remodeling of the cell wall** and hydrolytic activity related to cell wall. The plants are preparing the machinery to regulate the **starch and sugar metabolism** as suggested by the presence of beta amylase that degrade starch and starch synthase and trehalose-phosphate phosphate like. It is observed an enhancement of transcript related to amino sugar and nucleotide sugar metabolism. Also, the plants showed an increase in transcripts for the **respiration process** in which carbohydrates, fatty acid and amino acids are broken down to produce energy. The hydrolysis of glucose, fatty acids and proteins resulted in acetylcoa that enters in the TCA cycle.

Transporting activity is also enhanced as showed by the over-representation of amino acid transporters. This could suggest the necessity for an efficient partitioning of amino acids or peptides within the plant in the source-sink communication (Kohl et al., 2012). In addition, plants are responding with an increase in the transporting activities of macronutrients (such as nitrogen and

potassium) and micronutrients, (including zinc and iron transporter). It has been shown that a fine management of nutrients is very helpful to protect plants from temperature stress (high or low). For example, nitrogen in the form of nitric oxide (NO) is a highly reactive, membrane permanent free radical that has been shown to protect plants against stress by acting as an antioxidant directly scavenging the reactive oxygen species (ROS) generated under high or low temperatures. Some early report revealed that NO act as a signal in inducement of thermotolerance in plants, by activating active oxygen scavenging enzymes. In addition, Uchida et al. (2002) reported that NO protects the chloroplast against oxidative damage under heat stress by inducing expression of a gene encoding for the small heat shock protein 26 (HSP26). Also, the potassium is essential for many physiological processes such as photosynthesis, translocation of photosynthates into sink organs, maintenance of turgidity and activation of enzymes under stress conditions (Waraich et al., 2012).

Deep night. In the night, deep plants (refer to Table 5.5) are probably preparing the machinery for the synthesis of the **proteins involved** the cell cycle. In addition to transcripts possibly involved in the **regulation of cell cycle**, the cell are also organizing their cytoskeleton, as evident from the over-expression of **structural component of the cytoskeleton**. Deep plants prefer to enter the cell cycle in the nigh in comparison to the day, as evident also in the comparison "night - shallow-vs-deep", where the deep plants over-expressed transcripts involved in the cell cycle. This preference could be due to a fine regulation possibly related to the circadian clock and to the starch turnover. For example, *Arabidopsis thaliana* partitions a fraction of assimilates carbon into storage compounds in leaves to

support respiration and continued growth during the night, when photosynthesis in not possible. During the night, starch is degraded with a near-linear rate such that the starch reserves are almost completely utilized by dawn. This match between the length of time taken to degrade starch reserves and the length of the night is vitally important for normal plant growth (Graf et al., 2010). This hypothesis could be true also in the case of *Posidonia*.

5.6-Conclusions

In this study, Illumina next-generation sequencing and *de novo* assembly were applied to elucidate the transcriptome of *P. oceanica* for the first time. In total, 53,001 contigs were obtained, of these 34,88% were successfully annotated and 7,47% were mapped into 293 KEGG pathways. The data generated enormously increased the genomic resources in *P. oceanica* available so far and provided a global view of the *P. oceanica* transcriptome.

In addition, in this work, gene expression profiles of *Posidonia oceanica* growing at different depths, where light and temperature can be identified as the main diversifying forcing-factors, were described. The comparison between plants living in different conditions of light and temperature mainly highlighted the oxidative stress imposed in plants living in shallow waters by high light and high temperature. Plants responded by activating mechanism at multiple levels of organization (molecular, anatomical, morphological and metabolic), by adjusting the membrane systems and the cell wall architecture, by altering the cell cycle and the rate of cell division, and by metabolic tuning to provide energy for all these processes. Deep and shallow plants differ in the daily timing of the cell cycle, with deep plants preferring to grow in the night while shallow plants in the day.

The transcriptome data generated from this study also will represents a very valuable resource that could be used for subsequent investigations of the adaptation of *Posidonia oceanica*. It will be possible to better characterize the main pathways involved in the response to light and temperature i.e. to study the oxidative stress. The transcriptome also will facilitate genetic diversity analysis

and genome scan approaches, with the development of new EST-linked molecular markers.

5.7-References

Alcoverro, T., Manzanera, M., Romero, J. (1998). Seasonal and agedependent variability of *Posidonia oceanica* (L.) Delile photosynthetic parameters. *J. Exp. Mar. Biol. Ecol.*. 230: 1-13.

Ashburner, M., Ball, C. A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet.*, 25: 25-29.

Atkin, O. K. and M. G. Tjoelker. (2003). Thermal acclimation and the dynamic response of plant respiration to temperature, *Trends in Plant Science,* 8(7): 343-351.

Beckett B. S. (1986). Biology: Sensitivity and movement in plants. A modern introduction, (11): 122-130..

Birol, I., Jackman S.D., Nielsen, C.B., Qian, J.Q., Varhol, R., Stazyk, G., Morin, R. D., Zhao, Y., Hirst, M., Schein, J. E., Horsman, D. E., Connors, J. M., Gascoyne, R. D., Marra, M. A., Jones S. J. M. (2009). *De novo* transcriptome assembly with ABySS. *Bioinformatics,* 25: 2872-2877.

Borner, G. H. H., Sherrier, D. J., Stevens, T. J., Arkin, I. T., Dupree, P. (2002). Prediction of Glycosylphosphatidylinositol-Anchored Proteins in *Arabidopsis.* A Genomic Analysis 1, 129(June): 486-499.

Bulthuis, D. A. (1983). Effects of temperature on the photosynthesis- irradiance curve of the Australian seagrass, *Heterozostera tasmanica. Mar. Biol. Lett.*, 4: 47-57.

Bulthuis, D. A. (1987). Effects of temperature on photosynthesis and growth of seagrasses. *Aquat. Bot.*, 27: 27-40.

Casacuberta, J. M. and N. Santiago. (2003). Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. *Gene,* 311: 1-11.

Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., Yu, D. (2012). The role of WRKY transcription factors in plant abiotic stresses. *Biochimica et biophysica acta,* 1819(2): 120-8.

Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics,* 21: 3674-3676.

Dattolo, E., Gu, J., Mazzuca, S., Serra, I.A., Spadafora, A., Procaccini, G. (2013) *in press.* Transcriptional and proteomics profiles of Posidonia oceanica (seagrass) along its bathymetrical distribution. *Frontiers in Plant proteomics.*

Dennison, W. C. (1987). Effects of light on seagrass photosynthesis, growth and depth distribution. *Aquat. Bot.*, 27: 15-26.

Dennison, W. C., Orth, R. J., Moore, K. A., Stevenson, J. C., Carter, V., Kollar, S., Bergstrom, P. W., Batiuk, R. A. (1993). Assessing water quality with submersed aquatic vegetation. *Bioscience*, 43: 86–94.

Desikan, R, Burnett, E. C, Hancock, J. T., Neill, S. J. (1998) Harpin and hydrogen peroxide induce the expression of a homologue of gp91-phox in *Arabidopsis thaliana* suspension cultures. *J Exp Bot.*, 49: 1767–1771.

Desikan, R., Clarke, A., Hancock, J. T., Neill, S. J. (1999). H_2O_2 activates a MAP kinase-like enzyme in *Arabidopsis thaliana* suspension cultures. *J Exp Bot.*, 50: 1863–1866.

Desikan, R., Neill, S. J., Hancock, J. T. (2000). Hydrogen peroxide-induced gene expression in *Arabidopsis thaliana. Free Radic Biol Med.*, 28: 773–778.

Dodd, A. N., Salanthia, N., Hall, A., Ké, vei E., Tóoth, R., Nagy, F., Hibberd, J. M., Millar, A. J. (2005). Webb AAR: Plant circadian clocks increase photosynthesis, growth, survival and competitive advantage. *Science*, 309: 630-633.

Dos Reis, S. P., Lima, A. M., De Souza, C. R. B. (2012). Recent molecular advances on downstream plant responses to abiotic stress. *International journal of molecular sciences*, 13(7): 8628–47.

Duarte, C.M. (1991). Seagrass depth limits. *Aquat. Bot.*, 40: 363–377.

Dunton, K.H. (1990). Production ecology of *Ruppia maritima* L. s.l. and *Halodule wrightii* Aschers. in two subtropical estuaries. *J. Exp. Mar. Biol. Ecol.*, 143: 147–164.

Floris, M., Mahgoub, H., Lanet, E., Robaglia, C., Menand, B. (2009). Post-transcriptional regulation of gene expression in plants during abiotic stress. *Int. J. Mol. Sci.*, 10: 3168–3185.

Fraire-Velázquez S. and V. E. Balderas-Hernández. (2013). Abiotic Stress in Plants and Metabolic Responsesby in the book "Abiotic Stress - Plant Responses and Applications in Agriculture" edited by Kourosh Vahdati and Charles Leslie, ISBN 978-953-51-1024-8, InTech,

Franklin, F. A. and Whitelam, G. C. (2003). Light signals, phytochromes and cross-talk with other environmental cues. *Journal of Experimental Botany*, 55: 271- 276.

Franklin, K. A. (2009). Light and temperature signal crosstalk in plant development. *Current opinion in plant biology*, 12(1): 63–8.

Franssen, S. U., Gu, J., Bergmann N., Winters, G., Klostermeier, U. C., Rosenstiel, P., Bornberg-Bauer, E., Reusch, T. B. H. (2011) Transcriptomic resilience to global warming in the seagrass Zostera marina, a marine foundation species. *Proceedings of the Natural Academy of Sciences of the United States of America*, 108: 19276–19281.

Graf, A., Schlereth, A., Stitt, M., Smith, A. M. (2010). Circadian control of carbohydrate availability for growth in *Arabidopsis* plants at night. 107 (20): 9458–9463.

Grandbastien, M. A. (1998). Activation of plant retrotransposons under stress conditions. *Trends in Plant Science*, 3(5): 181-187.

Gray, W. M., Ostin, A, Sandberg, G., Romano, C. P., Estelle, M. (1998). High temperature promotes auxin-mediated hypocotyl elongation in *Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America,* 95(12): 7197–202.

Herzka, S.Z. and K. H. Dunton. (1997). Seasonal photosynthetic patterns of the seagrass *Thalassia testudinum* in the western Gulf of Mexico. *Mar. Ecol. Prog. Ser.,* 152: 103–117.

Hillman, K., Walker, D.I ., Larkum, A. W. D., McComb, A. J. (1989). Productivity and nutrient limitation. In: Larkum, A.W.D., McComb, A.J., Shepherd, S.A. (Eds.), Biology of Seagrasses: A Treatise on the Biology of Seagrasses with Special Reference to the Australian Region. Elsevier, Amsterdam, pp. 635–685.

Ito, H., Yoshida, T., Tsukahara, S., Kawabe, A. (2013). Evolution of the ONSEN retrotransposon family activated upon heat stress in Brassicaceae. *Gene,* 518(2): 256–61.

Jain, M. (2012). Next-generation sequencing technologies for gene expression profiling in plants. *Briefings in functional genomics,* 11(1): 63–70.

Kozaki, A. and G. Takeka. (1996). Photorespiration protects C3 plants from photooxidation. – *Nature,* 384: 557–560.

Kohl, S., Hollmann, J., Blattner, F. R., Radchuk, V., Andersch, F., Steuernagel, B., Schmutzer, T., Scholz, U., Krupinska, K., Weber, H., Weschke,W. (2012). A putative role for amino acid permeases in sink-source communication of barley tissues uncovered by RNA-seq. *BMC plant biology,* 12(1): 154.

Kovtun, Y., Chiu, W.L., Tena, G., Sheen, J. (2000). Functional analysis of oxidative stress-activated mitogen-activated protein kinase cascade in plants. *Proc Natl Acad Sci,* USA 97: 2940–2945.

Kumar, A, and J. L. Bennetzen. (2000). Retrotransposons: central players in the structure, evolution and function of plant genomes. *Trends in plant science,* 5(12): 509–10.

Lee, K.-S. and K. H. Dunton. (1996) Production and carbon reserve dynamics of the seagrass Thalassia testudinum in Corpus Christi Bay, Texas, USA. Mar. Ecol. Prog. Ser., 143: 201–210.

Lee, K.-S., Park, S. R., Kim, Y. K. (2007). Effects of irradiance, temperature, and nutrients on growth dynamics of seagrasses: A review. *Journal of Experimental Marine Biology and Ecology,* 350: 144-175.

Luning, K. and M. J. Dring. (1979). Continuous underwater light measurement in the sublittoral region. *Helgol. Wiss. Meeressunters.,* 32: 403–424.

Macauley, J. M., Clark, J. R., Price, W. A. (1988). Seasonal changes in the standing crop and chlorophyll content of *Thalassia testudinum* Banks ex König and its epiphytes in the northern Gulf of Mexico. *Aquat. Bot.,* 31: 277–287.

Mansour, A. (2007) Epigenetic activation of genomic retrotransposons, 6(2): 99–107.

Martin, J. A. and Z. Wang. (2011) Next-generation transcriptome assembly. *Nature reviews. Genetics,* 12(10): 671-82.

Markager, S. and K. Sand-Jensen. (1992). Light requirements and depth zonation of marine macroalgae. *Mar. Ecol. Prog. Ser.,* 88: 83-92.

Marsh Jr., J. A., Dennison, W. C., Alberte, R. S. (1986). Effects of temperature on photosynthesis and respiration in eelgrass (*Zostera marina* L.). *J. Exp. Mar. Biol. Ecol.,* 101: 257-267.

Masini, R.J. and C. R. Manning. (1997). The photosynthetic responses to irradiance and temperature of four meadow-forming seagrasses. *Aquat. Bot.,* 58: 21-36.

Mazzuca, S., Filadoro, D., Vannini, C., Marsoni, M., Cozza, R., Bracale, M., Pangaro, T., Innocenti , A. M. (2009) Seagrass light acclimation: 2-DE protein analysis in *Posidonia* leaves grown in chronic low light conditions. *Journal of Experimental Marine Biology and Ecology,* 374: 113-122.

Mazzuca, S., Björk, M., Beer, S., Felisberto, P., Gobert, S., Procaccini, G., Runcie, J., Silva, J., Borges, A.V., Brunet, C., Buapet, P., Champenois, W., Costa, M. M., D'Esposito, D., Gullström, M., Lejeune, P., Lepoint, G., Olivé, I., Rasmusson, L. M., Richir, J., Ruocco, M., Serra, I. A., Spadafora, A., Santos, R. (2013). Establishing research strategies, methodologies and technologies to link genomics and proteomics to seagrass productivity, community metabolism, and ecosystem carbon fluxes. *Frontiers in Plant Science,* 4 (38): 1-19.

McClintock, B. (1984). The significance of responses of the genome to challenge. *Science,* 226: 792-801.

Mirouze, M. and J. Paszkowski. (2011). Epigenetic contribution to stress adaptation in plants. *Current opinion in plant biology,* 14(3): 267-74.

Mittler, R. (2002). Oxidative stress, antioxidants and stress tolerance. *Trends Plant Sci,* 7:405-410.

Moore, K. A., Wetzel, R. L., Orth, R. J. (1997). Seasonal pulses of turbidity and their relations to eelgrass (*Zostera marina* L.) survival in an estuary. *J. Exp. Mar. Biol. Ecol.,* 215: 115-134.

Morelli, G. and I. Ruberti. (2002). Light and shade in the photocontrol of *Arabidopsis* growth. *Trends in plant science,* 7(9): 399-404.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 35: 182-185.

Olesen, B. and K. Sand-Jensen. (1993). Seasonal acclimatization of eelgrass *Zostera marina* growth to light. *Mar. Ecol. Prog. Ser.,* 94: 91-99.

Olesen, B., Enríquez, S., Duarte, C.M., Sand-Jensen, K. (2002). Depthacclimation of photosynthesis, morphology and demography of *Posidonia oceanica* and *Cymodocea nodosa* in the Spanish Mediterranean Sea. *Mar. Ecol. Prog. Ser.,* 236: 89-97.

Orth, R.J. and K. A. Moore. (1986). Seasonal and year-to-year variations in the growth of *Zostera marina* L. (eelgrass) in the lower Chesapeake Bay. *Aquat. Bot.,* 24: 335-341.

Penfield, S. (2008). Temperature perception and signal transduction in plants. *The New phytologist,* 179(3): 615–28.

Pérez, M. and J. Romero. (1992). Photosynthetic response to light and temperature of the seagrass *Cymodocea nodosa* and the prediction of its seasonality. *Aquat. Bot.,* 43: 51–62.

Phillips, R. C., McMillan, C., Bridges, K. W. (1983). Phenology of eelgrass, *Zostera marina* L., along latitudinal gradients in North America. *Aquat. Bot.,* 15: 145–156.

Procaccini, G., Dattolo, E., Innocenti, AM., Mazzuca, S., Cavallini, A., Natali, L.J., Bayer, P.G. (2010). Differential gene expression profiling under different light conditions in *Posidonia oceanica* (L.) Delile by SSH analysis. *Rapp Comm Int Mer Médit.,* 39: 641.

Procaccini, G., Beer, S., Björk, M., Olsen, J., Mazzuca, S., Santos,R. (2012). Seagrass ecophysiology meets ecological genomics: are we ready? *Mar. Ecol.,* 33: 522–527.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Ruan, Y. L. (2012) Signaling role of sucrose metabolism in development. *Molecular plant,* 5(4): 763–5.

Ruiz, J.M. and J. Romero. (2003). Effects of disturbances caused by coastal constructions on spatial structure, growth dynamics and photosynthesis of the seagrass *Posidonia oceanica.* Mar. *Pollut. Bull.,* 46: 1523–1533.

Ruocco, M., Brunet, C., Lorenti, M., Lauritano, C., D'Esposito, D., Riccio, M., Procaccini, G. (2012). *Posidonia oceanica* photoadaptation to the depth gradient. *Biol. Mar. Mediterr.,* 19 (1): 63-64.

Sand-Jensen, K. (1988). Minimum light requirements for growth in *Ulva lactuca. Mar. Ecol. Prog. Ser.* 50: 187–193.

Schindelman, G., Morikami, A, Jung, J., Baskin, T. I., Carpita, N. C., Derbyshire, P., McCann, M. C., Benfey, P. N. . (2001). COBRA encodes a putative GPI-anchored protein, which is polarly localized and necessary for oriented cell expansion in *Arabidopsis. Genes & development,* 15(9): 1115–27.

Setchell,W. A. (1929). Morphological and phenological notes on *Zostera marina* L. Univ. *Calif. Publ. Bot.,* 14: 389–452.

Short, F.T. and S. Wyllie-Echeverria. (1996). Natural and human-induced disturbance of seagrasses. *Environ. Conserv.,* 23: 17–27.

Sonenberg, N. and A. G. Hinnebusch. (2009). Regulation of translation initiation in eukaryotes: Mechanisms and biological targets. *Cell,* 136: 731–745.

Staiger, C. J. (2000). Signaling To the Actin Cytoskeleton in Plants. *Annual review of plant physiology and plant molecular biology,* 51: 257–288.

Strickland, J.D.H., 1958. Solar radiation penetrating the ocean. A review of requirements, data and methods of measurement, with particular reference to photosynthetic productivity. *J. Fish. Res. Board Can.*, 15: 453–493.

Suzuki, N., Miller, G., Morales, J., Shulaev, V., Torres, M. A., Mittler, R. (2011). Respiratory burst oxidases: the engines of ROS signaling. *Current opinion in plant biology*, 14(6): 691–9.

Thom, R. M. (1990). Spatial and temporal patterns in plant standing stock and primary production in a temperate seagrass system. *Bot. Mar.*, 33: 497–510.

Trapnell, C., Pachter, L., Salzberg, S. L. (2010). Transcript assembly and quantifi cation by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, 28: 511 - 515.

Tutin, T. G. (1942). *Zostera. J. Ecol.*, 30: 217–266.

Uchida, A.T., Jagendorf, T., Hibino. (2002). Effects of hydrogen peroxide and nitric oxide on both salt and heat stress tolerance in rice. *Plant Sci.* 163: 515-523.

Vermaat, J. E., Hootsmans, M. J. M., Nienhuis, P. H. (1987). Seasonal dynamics and leaf growth of *Zostera noltii Hornem.*, a perennial intertidal seagrass. *Aquat. Bot.*, 28: 287–299.

Vukich, M., Giordani, T., Natali, L., Cavallini, A. (2009). Copia and Gypsy retrotransposons activity in sunflower (Helianthus annuus L.). *BMC plant biology*, 9: 150.

Waraich, E. A., Ahmad, R., Halim, A., Aziz, T. (2012). Alleviation of temperature stress by nutrient management in crop plants: a review. *J. Soil Sci. Plant Nutr.*, 12(2): 221–244.

West, R.J. (1990). Depth-related structural and morphological variations in an Australian *Posidonia* seagrass bed. *Aquat. Bot.*, 36: 153–166.

Whitelam, G.C. and K. J., Halliday. (Eds) (2007). Light and Plant Development. Blackwell Publishing

Wissler, L., Dattolo, E., Moore, a D., Reusch, T. B. H., Olsen, J. L., Migliaccio, M., Bornberg-Bauer, E., Procaccini, G. (2009). Dr. Zompo: an online data repository for Zostera marina and *Posidonia oceanica* ESTs. *Database : the journal of biological databases and curation*, bap009. doi:10.1093/database/bap009

Wu, S., Li, R. W., Li, W., Li, C. (2012). Transcriptome Characterization by RNA-seq Unravels the mechanisms of butyrate-induced epigenomic regulation in bovine cells. *PLOS ONE* 7: e36940.

Zieman, J.C. and R. G. Wetzel. (1980). Productivity in seagrasses: methods and rates. In: Phillips, R.C., McRoy, C.P. (Eds.), Handbook of Seagrass Biology: An Ecosystem Perspective. Garland STPM Press, New York, pp. 87–116.

5.8-Appendix

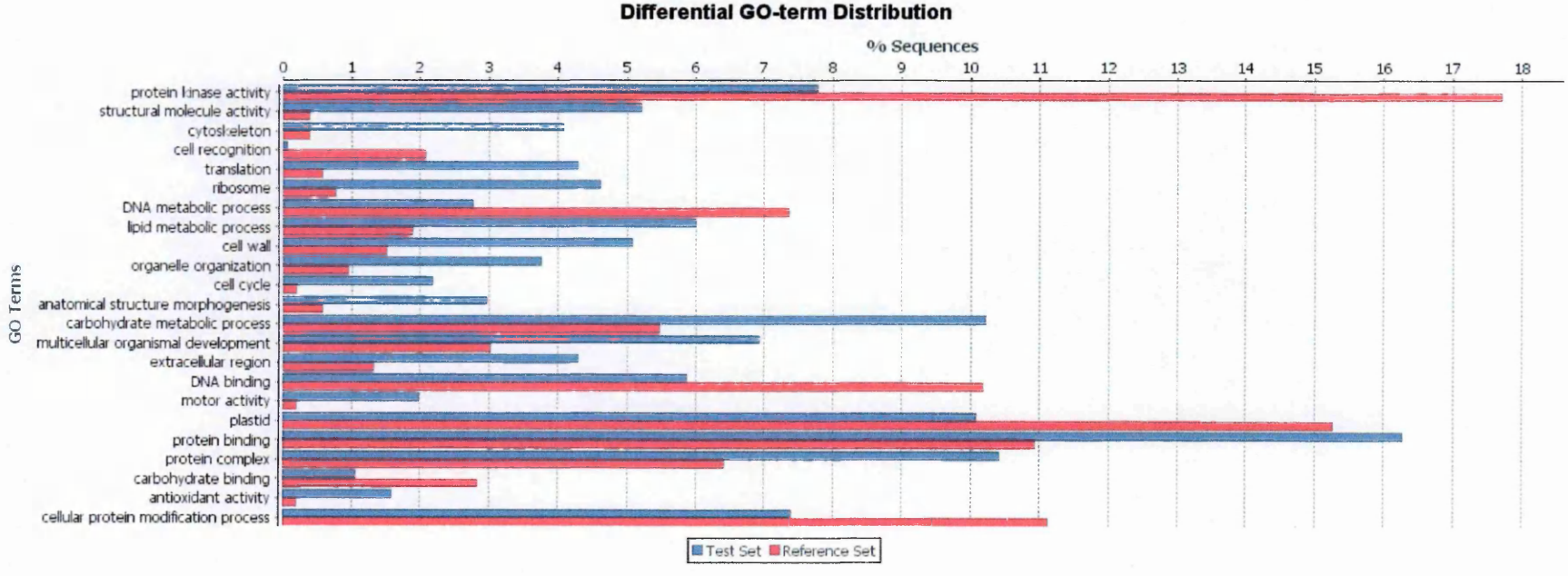


Figure 5.A1. Gene Ontology enrichment analysis. Test set = day shallow. Reference set = day deep.

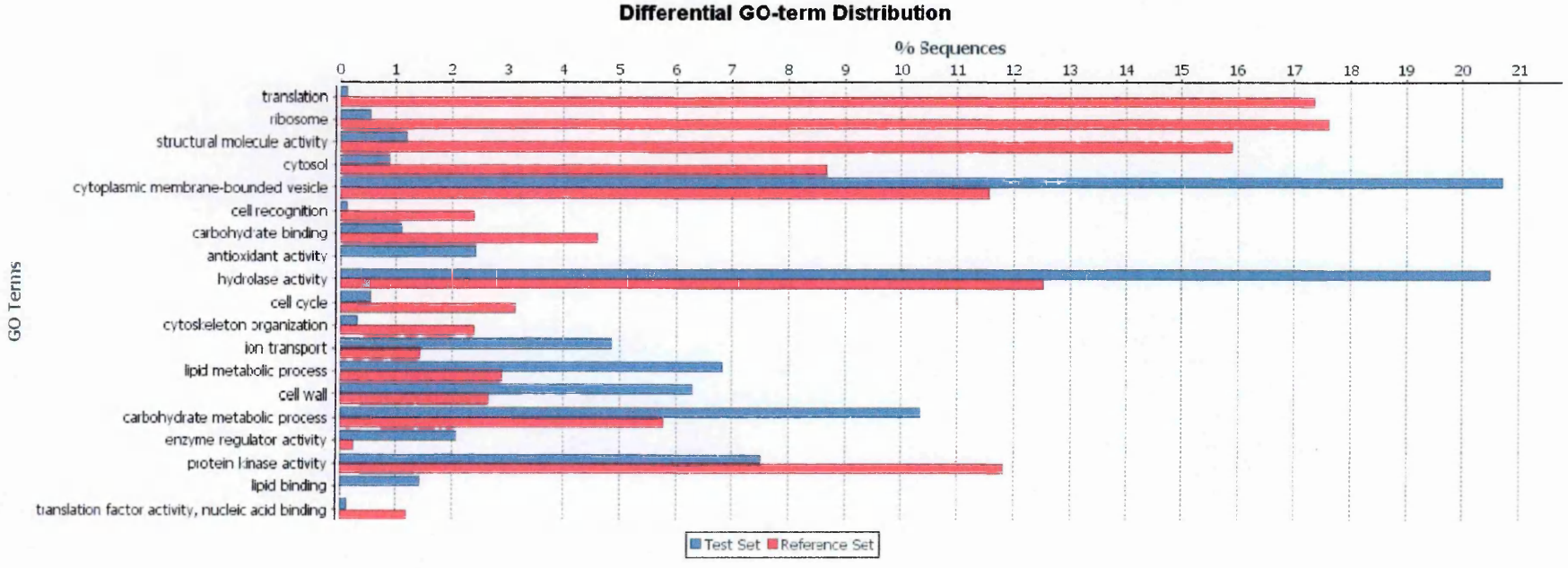


Figure 5.A2. Gene Ontology enrichment analysis. Test set = night shallow. Reference set = night deep.

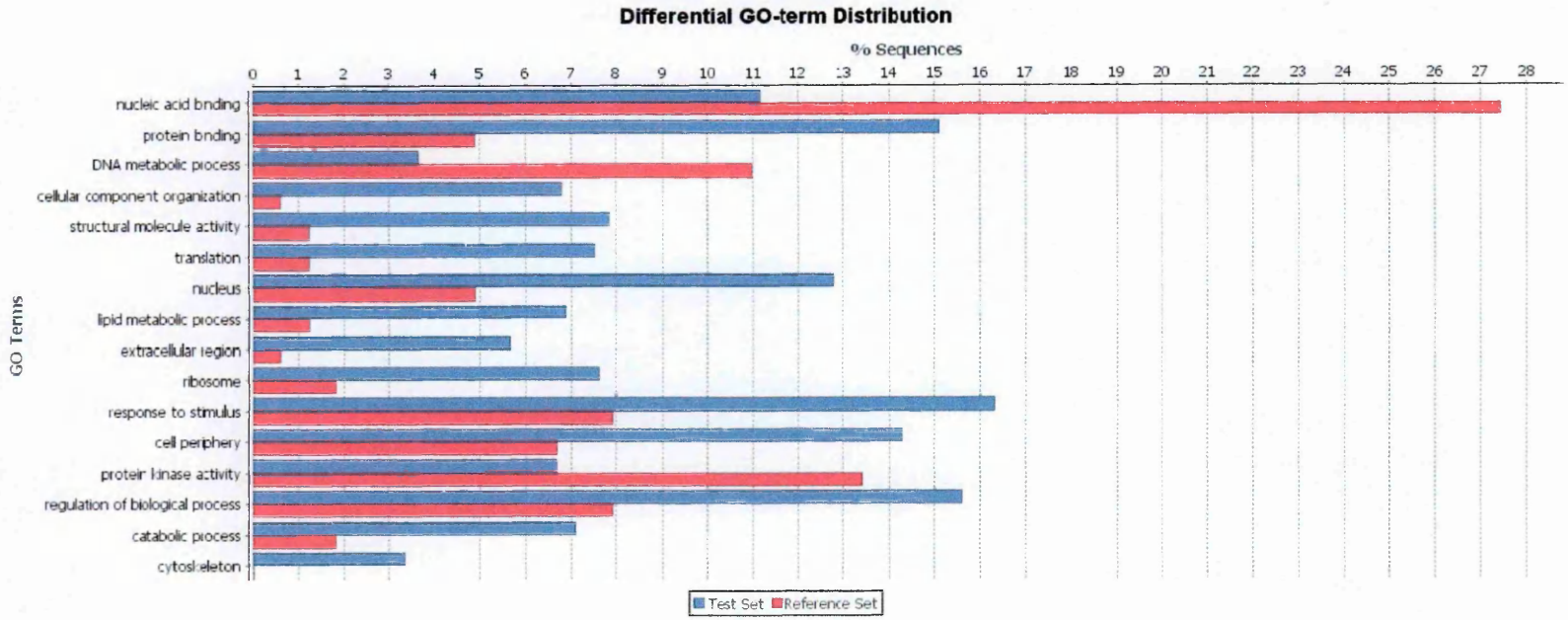


Figure 5.A3. Gene Ontology enrichment analysis. Test set = day shallow. Reference set = night shallow.

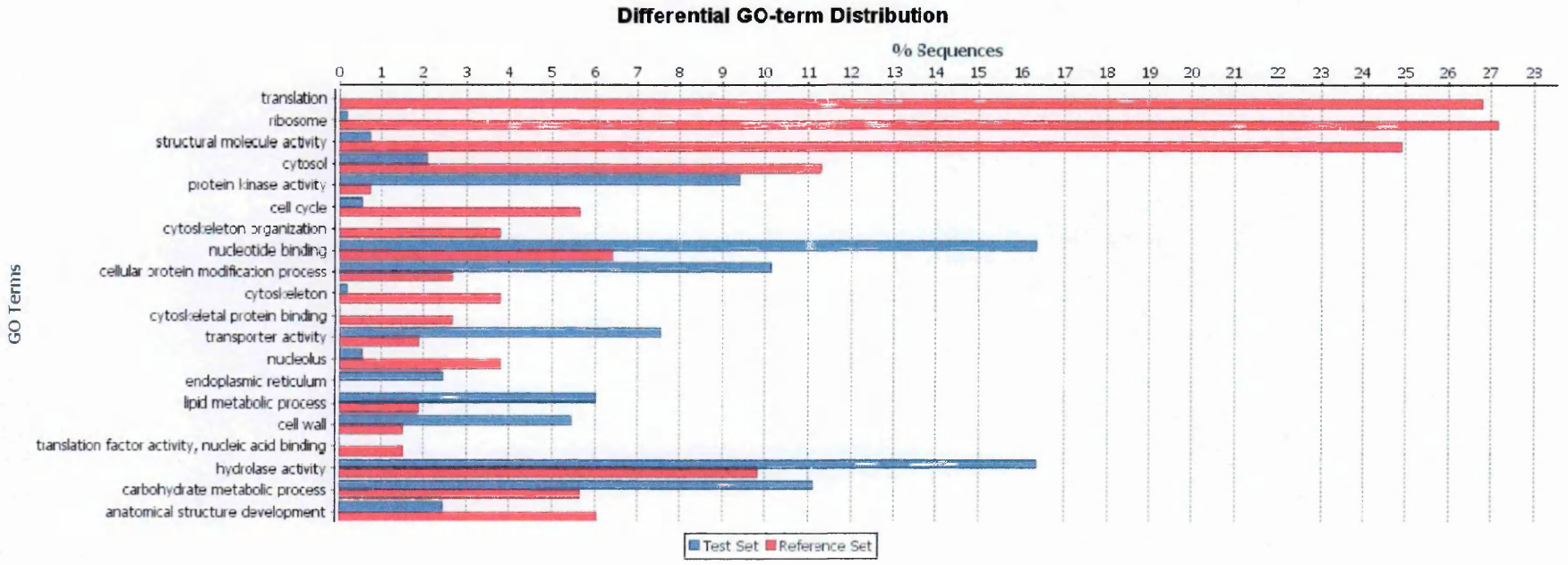


Figure 5.A4. Gene Ontology enrichment analysis. Test set = day deep. Reference set = night deep.

Chapter 6

6-Thesis conclusions

6.1-Main findings

The identification of the action of natural selection and the analysis of gene expression profiles in *P. oceanica* population living in different conditions of light and temperature allowed to achieve the objectives of this study and to have major knowledge about the plant response to environmental changes.

Objective 1-Detection of candidate genes under selection related to light and temperature changes. The development of EST-linked microsatellites (E-SSRs, Chapter 2) and the comparison of the performance of these markers with anonymous microsatellites (A-SSRs, Chapter 3) revealed their minor power in the population genetic analyses. Nevertheless, these markers, as linked to know function genes, have proved very useful for the identification of adaptive response to environmental changes, by means of the genome scan approach (Chapter 4). The genome scan performed along the depth gradient, detected no loci as outliers probably due to technical limitation related to the low number of loci used. Along the latitudinal gradient, 3 loci were identified to be under selection. The different results obtained along the two environmental gradients, where light and temperature represent the critical factors, may suggest that the signature of selection that has been detected, underlies adaptation to other environmental challenges than temperature and/or light alone. Alternatively, phenotypic plasticity is playing a stronger role than genetic adaptation in response to these environmental changes.

Objective 2-Increase of genomic resources for *P. oceanica* and identification of biological processes whose expression pattern is affected in different condition of light and temperature. The transcriptome sequencing, with the generation of a large quantity of sequences (53,001 contigs) and the successive analysis of differential gene expression highlighted the presence of oxidative stress imposed in plants living in shallow waters by the high light and high temperature and the plant response by activating multiple processes including the adjustments of the membrane systems and the cell wall architecture, the regulation of cell cycle and a metabolic tuning to provide energy for all these processes.

6.2-Future perspectives

The identification of adaptive genes by means of genome scan (Objective 1) represents the first step toward functional studies that will help to understand their potential effect on the phenotype and ultimately on the fitness (Dalziel et al., 2009). The analysis of sequences of these genes will elucidate the presence of variation at genetic levels and will allow a prediction about their impact. For example, there are many types of genetic variation that can affect the function and/or amount of proteins, including single nucleotide polymorphisms (SNPs), small insertion–deletion polymorphisms (indels), microsatellites and larger copy number variants (Feder, 2007). All of these types of genetic variation have been shown to contribute to variation in ecologically relevant traits (e.g. Hammock and Young 2004; Perry et al. 2007; Aminetzach et al. 2005; Schlenke and Begun 2004). Alternatively, there is the need to consider also the possibility that microsatellites themselves can be under selection and that the variation in the length of repeated

sequence may have important functional consequences. The variation in repeat units of SSRs present in 5'-UTR affects gene transcription and/or translation; in the coding region the variation of the sequence length can inactivate or activate genes or truncate protein; finally, SSRs present in 3'-UTR might be responsible for gene silencing or transcription slippage (Varshney et al., 2005). Recent studies have identified functional microsatellites that affect the individual fitness (Kashi and King, 2006; Gemayel et al., 2010). Variable tandem repeats in promoters can affect gene expression by altering the number of transcription factors binding sites, but can also induce changes in spacing between critical promoter elements. Tandem repeats in introns can influence gene expression by modulating the activity of RNA-binding proteins (Gemayel et al., 2012). Lastly, the genome scan does not give information of an association of molecular markers with ecological factors per se. Thus, a correlation of allele frequencies at outlier loci with environmental data will be necessary to identify those ecological factors potentially acting as selective force (Holderegger et al., 2008).

The transcriptome data generated from this study (Objective 2) increased the genomic resources available so far in *P. oceanica* and represent a very useful means that could be used for subsequent characterization of the main pathways involved in the response to light and temperature. The available transcriptome will also facilitate genetic diversity analysis with the development of EST-linked molecular markers.

6.3-References

Aminetzach, Y. T, Macpherson, J. M., Petrov, D. A. (2005). Pesticide resistance via transposition-mediated adaptive gene truncation in *Drosophila. Science*, 309: 764–767.

Dalziel, A. C, Rogers, S. M, Schulte, P. M. (2009). Linking genotypes to phenotypes and fitness: how mechanistic biology can inform molecular ecology. *Molecular Ecology*, 18: 4997–5017.

Feder, M. E. (2007). Evolvability of physiological and biochemical traits: evolutionary mechanisms including and beyond single-nucleotide mutation. *Journal of Experimental Biology*, 210: 1653–1660.

Hammock, E.A.D., and L. J., Young (2004). Functional microsatellite polymorphism associated with divergent social structure in vole species. *Molecular Biology and Evolution*, 21: 1057– 1063.

Holderegger, R., Herrmann, D., Poncet, B. et al. (2008). Land ahead: using genome scans to identify molecular markers of adaptive relevance. *Plant Ecology & Diversity*, 1: 273–283.

Gemayel, R.; Vinces, M.D.; Legendre, M.; Verstrepen, K.J. (2010). Variable Tandem Repeats Accelerate Evolution of Coding and Regulatory Sequences. *Annu. Rev. Genet.*, 44: 445-477.

Gemayel, R., Cho, J., Boeynaems, S., Kevin, J. V. (2012). Beyond Junk-Variable Tandem Repeats as Facilitators of Rapid Evolution of Regulatory and Coding Sequences. *Genes*, 3: 461-480; doi:10.3390/genes3030461

Kashi, Y., and D.G., King. (2006). Simple sequence repeats as advantageous mutators in evolution. *Trends in Genetics*, 22: 253-259.

Perry, G. H., Dominy, N. J., Claw, K. G. et al. (2007). Diet and the evolution of human amylase gene copy number variation. *Nature Genetics*, 39: 1256–1260.

Schlenke, T. A., and D. J. Begun. (2004). Strong selective sweep associated with a transposon insertion in *Drosophila simulans*. Proceedings of the National Academy of Sciences, USA, 101: 1626–1631.

Varshney, R. K., Graner, A., Sorrells, M. E. (2005). Genic microsatellite markers in plants: features and applications. *Trends in biotechnology*, 23(1): 48–55.

Acknowledgements

This thesis has seen through to completion with the support and encouragement of numerous people and I express my thanks to all those who contributed in many ways to the success of this study. In particular, I would like to express my gratitude to my Director of Study, Dr. Gabriele Procaccini. He provided the encouragement and advise necessary for me to proceed through the PhD program. Under his guidance I learned a lot.

I am also extremely grateful to my External Supervisor, Dr. Luisa Orsini, for her valuable help and her constructive criticism during these years.

Thanks to the member of "Procaccini's Group"for their friendship and for the good times spent together.

Lastly, I would to thank my family for their support throughout my research work.